THESIS / THÈSE

MASTER EN SCIENCES MATHÉMATIQUES

Dualité abstraite et théorie de composition de Robinson.

Hirsoux, Angélique

Award date:
1998

Facultés Universitaires Notre-Dame de la Paix
Namur
Faculté des Sciences - Département de Mathématiques

DUALITÉ ABSTRAITE ET THÉORIE DE COMPOSITION DE ROBINSON

Mémoire présenté pour l'obtention du grade
de Licencié en Sciences
Mathématiques
par

Promoteur NGUYEN V.H.

Angélique HIRSOUX

Année académique 1997-1998

REMERCIEMENTS

Je tiens tout d'abord à remercier mon professeur et promoteur V.H. NGUYEN, *qui malgré sa charge de doyen des Sciences, m'a accepté cette année comme mémorante et sans lequel ce mémoire n'aurait pas vu le jour. Ensuite,* A. HAMDI *pour ses précieux conseils le long de ce travail, quant à la rédaction de ce mémoire et pour son humour. Egalement, tous les professeurs et assistants qui m'ont suivis et vus évoluer au cours des années. Ma famille qui m'a soutenu durant toutes mes études. Michaël, pour la confiance qu'il a eu et qu'il a en moi. Et enfin, Christine, Denis, Laurent et Christel, Jean-François et Catherine, Alexandre et sa femme Karine, Ghislaine, Lissia et Lindsay, et tous les autres, compagnes et compagnons de route, avec lesquels j'ai partagé les meilleurs moments comme les plus difficiles durant toutes ces années d'étude.*

Résumé

Ce mémoire est dédié à l'étude d'une nouvelle notion de dualité dite *Dualité Abstraite*, impliquant une somme d'opérateurs qui unifie un grand nombre de résultats précédents de dualité classique, e.g. la dualité de Fenchel, et à son application à la théorie de composition d'opérateurs maximaux monotones selon Robinson.
Dans un premier temps, nous utilisons le principe de dualité abstraite appliqué à une somme d'opérateurs maximaux monotones afin de montrer que la somme est maximale monotone, et dans un second temps, nous montrons l'apport théorique et pratique de la composition : conservation de la monotonie et de la maximalité et réduction de la dimension initiale du problème variationnel.

Abstract

This report is dedicated to the study of a new notion of duality said *Abstract Duality*, implying a sum of operators, which unifies a very large number of previous results of classic duality, e.g. Fenchel duality, and to its application to the theory of Robinson's composition of maximal monotone operators.
In a first time, we use the principle of abstract duality to show that the sum of maximal monotone operators is maximal monotone, and in a second time, we show the theoretical and practical contribution of the composition : conservation of the maximal monotonicity properties and reduction of the initial variational problem dimension.

Table des matières

Introduction

L'analyse convexe est une branche importante des mathématiques et des mathématiques appliquées. Elle peut être considérée comme le champ dans lequel l'optimisation convexe et le calcul variationnel se développent; d'où l'intérêt considérable qu'a apporté un grand nombre de chercheurs à cette discipline.

L'analyse convexe offre d'une part une analyse géométrique et d'autre part une analyse complètement analytique. Depuis les années soixante, et encore aujourd'hui, elle reste une branche très active des mathématiques appliquées; l'une des raisons essentielles de cette vitalité est sans nul doute le nombre, la variété et l'importance de ses applications (voir [9, 10]).

Dans ce mémoire, nous nous intéressons à deux classes de problèmes : les problèmes variationnels que nous représentons ici par la recherche de zéro d'une somme de deux opérateurs maximaux monotones,

$$(PV1) \qquad Trouver \;\; x \;\; tel \;\; que \;\; 0 \in A(x) + B(x)$$

où A et B sont des opérateurs maximaux monotones, et les problèmes d'optimisation convexe qui s'expriment sous la forme suivante,

$$(P1) \qquad \begin{array}{l} \inf\{g(x) + h(Ax)\} \\ x \in I\!R^n \end{array}$$

La première partie de ce travail sera consacrée à la présentation d'un nouveau principe de dualité; dit *dualité abstraite* que nous appliquerons au problème (PV1). Ce principe a été récemment étudié par Attouch et Théra dans [1]. Nous en expliquons quelques avantages comme par exemple son utilisation pour résoudre une somme avec plus de deux opérateurs. Nous explicitons aussi une remarque importante amenée par l'application du nouveau principe, à savoir la différence d'écriture du résultat lorsque nous employons le principe de dualité abstraite. Ce principe proposé par Attouch et Théra unifie un grand nombre de principes de dualité comme la dualité classique de Fenchel et d'autres. Nous appliquons ce principe au problème de Fenchel-Rockafellar (P1) et nous vérifions que tous les résultats (connus et obtenus par le nouveau principe) sont bien identiques. Ce principe nous permet en particulier d'étendre nos connaissances au sujet de la somme d'opérateurs maximaux monotones et de montrer que, sous une condition plus faible que les précédentes, nous conservons la monotonie maximale pour l'opérateur-somme.

Dans la seconde partie, nous développons la notion de composition d'opérateurs et d'applications linéaires, selon Robinson [15]. Nous exposons également un autre principe de dualité qui étend le principe de dualité abstraite d'une somme d'opérateurs à une composition d'opérateurs, mais qui s'applique à un autre type de problème variationnel :

$$0 \in A(x) + QBP(x)$$

où A, B, P et Q sont, dans un premier temps, des multifonctions .

L'application de cet autre principe de dualité est motivante puisque elle nous permet de diminuer la dimension du problème initial. Nous présentons un exemple d'ordre économique où nous employons le principe, et nous diminuons clairement la dimension du problème à résoudre, voir [16]. L'intérêt que Robinson porte à cette composition est justifié, puisque ces compositions sont souvent présentes dans un grand éventail de problèmes. Certains, comme Fukushima [7], tentent de résoudre des problèmes où de telles compositions apparaissent, et pour la résolution ont besoin de savoir si la composition est maximale monotone. C'est pour cette raison que Robinson [15] a développé une théorie pour obtenir la monotonie maximale de cette composition. Nous reprenons et exposons les étapes de la recherche de Robinson à la fin de cette seconde partie.

Chapitre 1

Eléments d'Analyse Convexe

Comme la plupart des notions que nous utilisons dans ce travail émanent de cette partie, nous jugeons nécessaire de rappeler les notions que nous allons manipuler. Nous les préciserons dans certaines mesures suivant l'intérêt que nous y avons porté lors de l'élaboration de ce travail.

1.1 Fonctions convexes

Lorsque nous prononçons le mot « fonction convexe », nous pensons principalement à cette inégalité

$$f((\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y) \quad \forall x, y \in I\!R^n \quad \forall \lambda \in [0,1]$$

mais ce n'est pas toujours le cas car les valeurs $+\infty$ et $-\infty$ sont permises pour la fonction convexe f, et donc il peut arriver que nous trouvions des valeurs indéterminées du type $+\infty - \infty$. Nous devons donc définir les fonctions convexes de $I\!R^n$ vers $\overline{I\!R}$ autrement.

Commençons par définir une partie convexe :

Définition 1.1.1 *Une partie non vide $C \subseteq I\!R^n$ est dite* convexe *si et seulement si quels que soient x et y dans C, et quel que soit $\lambda \in [0,1]$, $\lambda x + (1-\lambda)y$ est dans C également, i.e. le segment de droite reliant deux éléments quelconques de C doit se trouver dans C.*

Nous pouvons donner quelques exemples évident de parties convexes : un segment de droite, une boule de $I\!R^n$ pour toute p-norme (avec $p > 1$), l'espace $I\!R^n$, l'ensemble vide, un cube,...

Définition 1.1.2 *Soit une partie non vide $C \subseteq I\!R^n$. L'*enveloppe convexe *de C, notée conv C, est l'intersection de toutes les parties convexes de $I\!R^n$ qui contiennent C.*

Avant de définir une fonction convexe, nous donnons une notion intermédiaire d'un ensemble associé à la fonction :

Définition 1.1.3 *Soit $f : I\!R^n \longrightarrow \overline{I\!R}$ une fonction pouvant prendre les valeurs $-\infty$ et $+\infty$ et ayant une partie convexe non vide $C \subseteq I\!R^n$ pour domaine. L'ensemble défini par :*

$$\{(x, \mu) \in C \times I\!R \mid \mu \geq f(x)\} \tag{1.1}$$

s'appelle épigraphe *de f et est noté epi f. C'est une partie de $I\!R^{n+1}$.*

Remarquons que dans la définition de epi f, seuls les nombres μ réels sont pris en compte; $\mu = \pm\infty$ ne sont pas des valeurs admises. Graphiquement, l'épigraphe de f est tout ce qui se trouve « au-dessus » du graphe de f.

Maintenant nous pouvons définir une fonction convexe admettant des valeurs infinies :

Définition 1.1.4 *La fonction f sera dite* convexe *si et seulement si l'épigraphe de f est une partie convexe de $I\!R^{n+1}$.*

Remarque 1.1.5 *La fonction $f : I\!R^n \longrightarrow \overline{I\!R}$ sera dite* concave *si et seulement si la fonction $(-f)$ est convexe, et elle sera dite* affine *si et seulement si elle est à la fois convexe et concave.*

Définition 1.1.6 *Le* domaine effectif *d'une fonction convexe f sur $C \subseteq I\!R^n$ est l'ensemble des points où f ne vaut pas $+\infty$:*

$$dom\ f \stackrel{déf}{=} \{x \in C \mid f(x) < +\infty\} \tag{1.2}$$

Cet ensemble est en fait la projection de epi f sur $I\!R^n$. C'est un convexe.

Définition 1.1.7 *Une fonction $f : I\!R^n \longrightarrow \overline{I\!R}$ sera dite* coercive *si*

$$\lim_{\|x\| \to +\infty} f(x) = +\infty$$

Définition 1.1.8 *Une fonction $f : I\!R^n \longrightarrow \overline{I\!R}$ sera dite* stricte *s'il existe un élément de $I\!R^n$ tel que l'image par f de cet élément est strictement inférieure à la valeur $+\infty$, i.e. $\exists\ x \in I\!R^n$ tel que $f(x) < +\infty$.*

La définition suivante étend la définition de fonction stricte, en demandant une condition supplémentaire dans sa définition :

Définition 1.1.9 *Une fonction $f : I\!R^n \longrightarrow \overline{I\!R}$ sera dite* propre *si et seulement si son épigraphe est non vide et ne contient aucune droite verticale. Cela signifie que :*

1. *$\exists\ x \in I\!R^n \mid f(x) < +\infty$, ou de manière équivalente, dom $f \neq \emptyset$,*

2. $\forall x \in \mathbb{R}^n \quad f(x) > -\infty$,

ou encore :
Une fonction est propre si elle ne prend pas la valeur $+\infty$ identiquement et si elle ne prend jamais la valeur $-\infty$.

Exemples :

- Toute fonction convexe f finie sur une partie convexe non vide C et prolongée hors de C en posant $f(x) \stackrel{\text{déf}}{=} +\infty \quad \forall x \notin C$ est convexe propre.
- La fonction $x \to \|x\|_2^2$ est convexe propre sur $\mathbb{R}^n$ où $\|x\|_2 = \sum_{i=1}^n x_i^2$.

Définition 1.1.10 *Une fonction $f : \mathbb{R}^n \longrightarrow \overline{\mathbb{R}}$ définie sur une partie $C \subseteq \mathbb{R}^n$ est dite* semi-continue inférieurement *au point $x_0 \in C$ si et seulement si :*

$$\forall \varepsilon > 0, \quad \exists\, \delta > 0 \ \ tel \ \ que \ \ f(x) > f(x_0) - \varepsilon$$

ce qui est équivalent à

$$\forall k < f(x_0), \quad \exists\, \mathcal{U} \ (voisinage \ \ de \ \ x_0) \ \ tel \ \ que \ \ f(x) > k \ \ \forall x \in \mathcal{U}$$

ou encore

$$\liminf_{x \to x_0} f(x) > f(x_0)$$

Une fonction est semi-continue inférieurement sur C si et seulement si la fonction est semi-continue inférieurement en chaque $x \in C$.

Revenons à la fonction affine, vue à la Remarque 1.1.5, et définissons-la plus précisément :

Soient $a \in \mathbb{R}^n$ et $b \in \mathbb{R}$.
Une fonction affine F est une fonction du type $F(x) \equiv a^T x - b : \mathbb{R}^n \longrightarrow \mathbb{R} \ \ \forall x \in \mathbb{R}^n$.
Remarquons que cette fonction est toujours à valeurs réelles, et rappelons qu'elle est à la fois convexe et concave.

Considérons toutes les fonctions affines minorantes F d'une fonction convexe f, i.e.

$$F(x) \leq f(x) \ \ \forall x \in \mathbb{R}^n$$

Nous notons par $\mathcal{A}(f)$ l'ensemble de toutes les fonctions affines minorantes de f.

Définition 1.1.11 *Soit $f : \mathbb{R}^n \longrightarrow \overline{\mathbb{R}}$ convexe et $\mathcal{A}(f)$ défini comme ci-dessus. La fermeture $\overline{f}$ de f s'exprime par :*

$$\overline{f}(x) = \sup_{F \in \mathcal{A}(f)} F(x) \tag{1.3}$$

$\overline{f}(x)$ est convexe.
La fonction f est dite *fermée* si et seulement si $f = \overline{f}$.

Théorème 1.1.12 *Soit f une fonction convexe, propre de $I\!R^n$ dans $\overline{I\!R}$. Les assertions suivantes sont équivalentes :*

(a) f est fermée, i.e. $f = \overline{f}$,

(b) L'épigraphe de f est fermé dans $I\!R^{n+1}$,

(c) $\{x \in I\!R^n \mid f(x) \leq \alpha\}$ est fermé pour tout $\alpha \in I\!R$,

(d) f est semi-continue inférieurement sur $I\!R^n$.

Le théorème suivant sera utile dans le développement du problème de Fenchel-Rockafellar.

Théorème 1.1.13 *Si $g : I\!R^m \longrightarrow I\!R^n$ est une fonction linéaire, si $f_1, f_2 : I\!R^n \longrightarrow \overline{I\!R}$ sont convexes et si α est un scalaire strictement positif :*

1. La composée $f_1 \circ g$, où $(f_1 \circ g)(x) \stackrel{déf}{=} f_1(g(x))$, est une fonction convexe.

2. La somme $f_1 + f_2$, où $(f_1 + f_2)(x) \stackrel{déf}{=} f_1(x) + f_2(x)$, est une fonction convexe.

3. La produit $\alpha\ f_1$, où $(\alpha\ f_1)(x) \stackrel{déf}{=} \alpha\ f_1(x)$, est une fonction convexe.

Une dernière notion importante de cette section, que nous utiliserons par la suite, est la fonction indicatrice convexe :

Définition 1.1.14 *Etant donnée une partie convexe $C \subseteq I\!R^n$, la* fonction indicatrice *de C notée $I_C(\cdot)$ est définie par :*

$$I_C(x) = \begin{cases} 0 & si \quad x \in C \\ +\infty & si \quad x \notin C \end{cases}$$

1.2 Intérieur relatif

Définition 1.2.1 *Etant donnée une partie convexe $C \subseteq I\!R^n$, l'* enveloppe affine *de C est le plus petit sous-espace affine de $I\!R^n$ qui contient C. En d'autres termes, c'est l'intersection de tous les sous-espaces affines de $I\!R^n$ qui contiennent C. Elle est notée aff C.*

Exemples
Dans $I\!R^3$, l'enveloppe affine d'une boule est $I\!R^3$. L'enveloppe affine d'un point est ce point, alors que l'enveloppe affine d'un segment de droite est la droite affine supportée par ce segment.

Notons par B, la boule unité avec la norme Euclidienne dans $\mathbb{R}^n$, que nous définissons comme suit :

$$B = \{x \mid \|x\| \leq 1\} = \{x \mid d(x,0) \leq 1\}$$

Définition 1.2.2 *L'*intérieur *de C, noté $Int\ C$, est l'ensemble :*

$$Int\ C = \{x \mid \exists\ \varepsilon > 0,\ x + \varepsilon B \subset C\}$$

Dans le cas des ensembles convexes, la notion d'intérieur peut être englobée dans une notion plus commode d'intérieur relatif.

Définition 1.2.3 *L'* intérieur relatif *d'une partie convexe $C \subseteq \mathbb{R}^n$, noté $ri\ C$, est l'intérieur de C lorsque C est considéré comme une partie de son enveloppe affine $aff\ C$* [1]*. En d'autres termes :*

$$ri\ C = \{x \in aff\ C \mid \exists\ \varepsilon > 0\quad (x + \varepsilon B) \cap aff\ C \subseteq C\}$$

Exemples

- Dans $\mathbb{R}^n$, l'intérieur relatif d'un point est ce point.
- L'enveloppe affine d'une boule de $\mathbb{R}^3$ étant $\mathbb{R}^3$, son intérieur relatif coïncide avec son intérieur, et en général, si $C \subseteq \mathbb{R}^n$ est une partie convexe de dimension n, on a ri $C = \overset{\circ}{C}$.
- Dans $\mathbb{R}^2$, l'enveloppe affine du segment $\{(x,0) \mid 1 \leq x \leq 2\}$ est le premier axe de coordonnées $\{(x,0) \mid x \in \mathbb{R}\}$ et son intérieur relatif est son intérieur lorsque nous le considérons comme un intervalle de $\mathbb{R}$, c'est donc $\{(x,0) \mid 1 < x < 2\}$.

1.3 Fonctions conjuguées et dualité

Définition 1.3.1 *Soit $f : \mathbb{R}^n \longrightarrow \overline{\mathbb{R}}$ une fonction convexe, la fonction f^* définie par :*

$$\begin{array}{rccl} f^*: & \mathbb{R}^n & \longrightarrow & \overline{\mathbb{R}} \\ & x^* & \longrightarrow & \sup\{\langle x, x^*\rangle - f(x) \mid x \in \mathbb{R}^n\} \end{array}$$

est appelée conjuguée *de f. L'opération, qui à une fonction convexe f associe sa conjuguée f^* (i.e. $f \rightsquigarrow f^*$), est appelée* transformation de Fenchel.

Propriétés de la conjuguée

1. La fonction f^* est convexe fermée.

1. «ri» signifie *relative interior*

2. La conjuguée $(f^*)^*$ de f^*, que nous notons f^{**} est égale à la fermeture $\overline{f}$ de f. Si la fonction f est fermée, la transformation de Fenchel est donc involutive, c'est-à-dire que si nous l'appliquons deux fois, nous retrouvons la fonction de départ.

3. Les fonctions constantes $+\infty$ et $-\infty$ sont conjuguées l'une de l'autre. Comme ce sont les seules fonctions convexes impropres et fermées, tous les autres couples (f, f^*) sont donc constitués de fonctions *propres.*

4. Pour calculer la conjuguée f^* de f, nous pouvons nous restreindre à l'intérieur relatif du domaine de f,
$$f^*(x^*) = \sup\{\langle x, x^*\rangle - f(x) \mid x \in \text{ri dom } f\}$$
(pour la vérification de ce résultat, voir [17])

5. Nous avons toujours
$$\begin{aligned} f^*(x^*) &= \sup_{x\in I\!R^n}\{\langle x, x^*\rangle - f(x)\} \\ &\geq \; < x, x^* > -f(x) \qquad \forall x^* \end{aligned}$$
i.e.
$$\langle x, x^*\rangle \leq f(x) + f^*(x^*) \;\; \forall x, x^* \tag{1.4}$$
Cette relation est connue sous le nom d'*inégalité de Young-Fenchel.*

6. Les couples (x, x^*) pour lesquels l'inégalité de Young-Fenchel (1.4) est en fait une égalité forment un ensemble qui peut être considéré comme le graphe d'un opérateur dit *sous-différentiel* de f (noté ∂f). Nous reparlerons de ces notions aux sections 1.5 et 1.6.

Proposition 1.3.2 *Soit une fonction convexe f sur $I\!R^n$. Alors,*

- *si $\alpha \in I\!R \setminus \{0\}$ et $f_0(x) = f(\alpha x)$ $\forall x$, alors $f_0^*(y^*) = f^*(y^*/\alpha)$ $\forall y^*$,*
- *si $\alpha \in I\!R^+ \setminus \{0\}$, alors $(\alpha f)^*(y^*) = \alpha f^*(y^*/\alpha)$ $\forall y^*$.*

En particulier, si $\alpha = -1$, alors $(-f)^*(y^*) = -f^*(-y^*)$ $\forall y^*$.

Considérons le problème suivant
$$\begin{gathered}\inf f(x) \\ x \in I\!R^n\end{gathered}$$
que nous appelons problème primal.

Théorème 1.3.3 *Soit une fonction convexe, propre f sur $I\!R^n$. Alors, nous avons*
$$\sup_{y^*\in I\!R^n}\{-f^*(y^*)\} \leq \inf_{x\in I\!R^n} f(x)$$
i.e. les solutions du problème primal sont supérieures aux solutions données par la résolution du problème de la partie gauche de l'inégalité.

1.4 Cône normal

Définition 1.4.1 *Soit C une partie de $I\!R^n$ et un vecteur $h \in I\!R^n$. Nous dirons que le vecteur h est* tangent *à la partie C au point $\overline{x} \in C$ si et seulement si les trois conditions suivantes sont satisfaites :*

1. *Il existe une suite $(h_k)_{k \in I\!N}$ convergeant vers h.*
2. *Il existe une suite $(t_k)_{k \in I\!N}$ de réels strictement positifs convergeant vers 0.*
3. *Ces suites sont telles que $\overline{x} + t_k h_k \in C \quad \forall k \in I\!N$.*

Définition 1.4.2 *Soit C une partie de $I\!R^n$ et $\overline{x} \in C$. L'ensemble de tous les vecteurs h tangents à C au point $\overline{x}$ est appelé* cône tangent à C en $\overline{x}$ *et est noté $T_C(\overline{x})$.*

Exemples

1. Si $C = I\!R^n$ et si $\overline{x} \in C$, alors tout vecteur h est tangent à C en $\overline{x}$ car il suffit de prendre dans la défintion $h_k \stackrel{déf}{=} h \; \forall k$ et $t_k \stackrel{déf}{=} \frac{1}{k} \; \forall k \geq 1$. Nous avons donc $T_{I\!R^n}(\overline{x}) = I\!R^n$.
2. Si $C = \{0\}$, alors le seul $\overline{x}$ possible est $\overline{x} = 0$ et $h = 0$ est le seul vecteur tangent.

Le cas convexe

Définition 1.4.3 *Soit une partie convexe non vide $C \subseteq I\!R^n$ et $\overline{x} \in C$. Un vecteur non nul $d \in I\!R^n$ est une* direction admissible *pour C à partir de $\overline{x}$ si et seulement si :*

$$\exists \, \varepsilon > 0 \quad \forall t \in [0, \varepsilon] \qquad \overline{x} + td \in C$$

Cela signifie que le segment $\{\overline{x} + td \mid t \in [0, \varepsilon]\}$ est contenu dans C. Autrement dit, il faut pouvoir suivre la direction d « pendant un certain temps » (représenté par ε) à partir du point $\overline{x}$ sans sortir de C. Notons l'ensemble de ces directions $\mathcal{A}dm_C(\overline{x})$.

Théorème 1.4.4 *Lorsque la partie C est convexe, le cône tangent à C en $\overline{x} \in C$ est la fermeture de l'ensemble des directions admissibles pour C à partir de $\overline{x}$:*

$$\begin{aligned} T_C(\overline{x}) &= \{0\} \bigcup \overline{\mathcal{A}dm_C(\overline{x})} \\ &= \{0\} \bigcup \{h \neq 0 \mid h = \lim_{k \to \infty} h_k, \; h_k \in \mathcal{A}dm_C(\overline{x}) \; \forall k\} \end{aligned}$$

Définition 1.4.5 *Un vecteur $v \in I\!R^n$ est dit* normal *à une partie C de $I\!R^n$ au point $\overline{x} \in C$ si et seulement si*

$$\langle v, h \rangle \leq 0 \qquad \forall h \in T_C(\overline{x})$$

Un vecteur normal doit donc faire un angle obtus ou droit avec tout vecteur tangent. Nous constatons que si v est normal à C en $\overline{x}$ et si $\lambda > 0$, alors λv est encore normal, et que le vecteur nul est toujours un vecteur normal.

Définition 1.4.6 *L'ensemble de tous les vecteurs v normaux à C en $\overline{x}$ s'appelle* cône normal à C en $\overline{x}$ *et est noté $N_C(\overline{x})$.*

Exemples

1. Si $C = I\!R^n$, nous savons que $T_C(\overline{x}) = I\!R^n$ et l'ensemble des vecteurs qui font un angle obtus ou droit avec tous les vecteurs de $I\!R^n$ ne contient qu'un seul élément; c'est le vecteur nul. Donc $N_{I\!R^n}(\overline{x}) = \{0\}$.

2. Si $C = \{0\}$, nous savons que le cône tangent à C est C lui-même, et comme tout vecteur fait un angle obtus ou droit avec le vecteur nul, on a $N_{\{0\}}(0) = I\!R^n$.

Nous voyons facilement d'après ces définitions que $N_C(\overline{x})$ est toujours un cône convexe fermé qui contient l'origine. Le cône tangent était lui aussi un cône contenant l'origine mais il n'est pas toujours convexe.

Théorème 1.4.7 *Lorsque la partie C est convexe, si $\overline{x} \in C$, le cône normal peut s'exprimer de la façon suivante:*

$$N_C(\overline{x}) = \{v \in I\!R^n \mid \langle v, x - \overline{x}\rangle \leq 0 \ \forall x \in C\} \tag{1.5}$$

A la Section 1.5, nous verrons encore une autre expression du cône normal à une partie convexe en termes d'opérateur.

Dans cette section, les notions développées sont issues du cours d'introduction à l'optimisation de première licence de M. STRODIOT J.-J.

1.5 Sous-gradients

Soit $f : I\!R^n \longrightarrow I\!R$ une fonction convexe mais *non nécessairement différentiable* et $X \subseteq I\!R^n$ une partie convexe non vide.
Cette particularité quant à la différentiabilité non nécessaire de f nous amènera à la notion d'opérateur multivoque, c'est-à-dire une fonction de source $I\!R^n$ et de but $\mathcal{P}(I\!R^n)$, où $\mathcal{P}(I\!R^n)$ désigne l'ensemble des parties de $I\!R^n$.
Nous verrons ces notions d'opérateur plus en détail dans la Section 1.6.

Lorsqu'une fonction $f : I\!R^n \longrightarrow I\!R$ est convexe, différentiable et de domaine $I\!R^n$, nous pouvons montrer que l'inégalité

$$f(z) \geq f(x) + \langle \nabla f(x), z - x\rangle$$

est vérifiée pour tout $z \in \mathbb{R}^n$.
Cette inégalité signifie que le graphe de la fonction affine

$$f(x) + \langle \nabla f(x), z - x \rangle$$

est un hyperplan passant par $(x, f(x))$, et qui est situé «en dessous» de l'épigraphe de f.

Nous aimerions obtenir une caractérisation semblable lorsque la fonction f n'est pas nécessairement différentiable. Donc, nous allons généraliser les notions de *gradient* et d'*application différentielle* pour des fonctions convexes générales. La généralisation du gradient sera appelée *sous-gradient* et celle d'application différentielle, *sous-différentiel.* En quelque sorte, un sous-gradient d'une fonction convexe en un point est un vecteur qui *aurait pu* être gradient de la fonction en ce point, si la fonction avait été différentiable.

Définition 1.5.1 *Soit $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ une fonction convexe. Un vecteur $x^* \in \mathbb{R}^n$ est un* sous-gradient *de f au point $x \in \mathbb{R}^n$ si et seulement si*

$$f(z) \geq f(x) + \langle x^*, z - x \rangle \quad \forall z \in \mathbb{R}^n$$

Un sous-gradient x^* de f en x détermine donc une fonction affine dont le graphe est un hyperplan qui supporte l'épigraphe de f en passant par $(x, f(x))$. De plus, dans le cas différentiable, nous voyons que le gradient est un sous-gradient.

Remarque 1.5.2 *Dans le cas où $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ est une fonction concave, nous avons qu'un vecteur $x^* \in \mathbb{R}^n$ est un* sous-gradient *de f au point $x \in \mathbb{R}^n$ si et seulement si*

$$f(z) \leq f(x) + \langle x^*, z - x \rangle \quad \forall z \in \mathbb{R}^n$$

Définition 1.5.3 *L'ensemble de tous les sous-gradients de f en $x \in \mathbb{R}^n$ s'appelle le* sous-différentiel de f en x *et est noté $\partial f(x)$. Nous avons donc*

$$\partial f(x) \stackrel{\text{déf}}{=} \{x^* \in \mathbb{R}^n \mid f(z) \geq f(x) + \langle x^*, z - x \rangle \quad \forall z \in \mathbb{R}^n\}$$

L'opérateur (éventuellement multivoque)

$$\begin{array}{rccc} \partial f : & \mathbb{R}^n & \longrightarrow & \mathcal{P}(\mathbb{R}^n) \\ & x & \longrightarrow & \partial f(x) \end{array}$$

est appelé sous-différentiel *de f.*

Dans le cas général, $\partial f(x)$ peut être vide, et s'il est non vide, il peut posséder un nombre infini d'éléments. Si $\partial f(x) \neq \emptyset$, nous dirons que f est sous-différentiable *au point x.*

Remarquons que $\partial f(x)$ est toujours un ensemble convexe fermé.

Un exemple particulièrement important de fonction convexe est la fonction *indicatrice* d'une partie convexe non vide C de $I\!R^n$:

$$I_C(\cdot) : I\!R^n \longrightarrow \overline{I\!R}$$

$$I_C(x) = \begin{cases} 0 & \text{si } x \in C \\ +\infty & \text{si } x \notin C \end{cases}$$

Alors nous avons

$$x^* \in \partial I_C(x) \Leftrightarrow I_C(z) \geq I_C(x) + \langle x^*, z - x \rangle \ \forall z \in I\!R^n$$

Cette inégalité est impossible si $x \notin C$, nous devons donc avoir que $I_C(x) = 0$ et elle doit seulement être vérifiée pour tout $z \in C$, car elle sera toujours vraie si $z \notin C$. Donc il nous reste :

$$\begin{aligned} x^* \in \partial I_C(x) \quad &\Leftrightarrow \quad \langle x^*, z - x \rangle \leq 0 \ \forall z \in C \\ &\Leftrightarrow \quad x^* \in N_C(x) \end{aligned}$$

où $N_C(x)$ est le *cône normal* à C au point $x \in C$.

Nous pouvons en effet démontrer que l'ensemble $\partial f(x)$ est réduit à un singleton $\{x^*\}$ si et seulement si f est finie au voisinage de x et est différentiable en x, et $\nabla f(x) = x^*$.

Théorème 1.5.4 *Soit $f : I\!R^n \longrightarrow \overline{I\!R}$ une fonction convexe, propre.*

1. *f n'est pas sous-différentiable hors de son domaine,* i.e.
   $\forall x \notin dom\ f$, nous avons $\partial f(x) = \emptyset$.
2. *f est toujours sous-différentiable sur l'intérieur relatif de son domaine, i.e*
   $\forall x \in ri\ dom f$, $\partial f(x) \neq \emptyset$.
3. *$\partial f(x)$ est un ensemble borné et non vide si et seulement si $x \in (\overset{\circ}{dom}\ f)$.*

Pour la troisième partie de ce théorème, rappelons que l'ensemble $\partial f(x)$ est toujours fermé et convexe. S'il est borné, il sera donc compact.
Le résultat très important à retenir de ce théorème est que toute fonction convexe propre est sous-différentiable sur l'intérieur relatif de son domaine.

Théorème 1.5.5 *Soit $f : I\!R^n \longrightarrow \overline{I\!R}$ une fonction convexe propre et soit $x \in I\!R^n$. Les assertions suivantes sont équivalentes :*

(a) $x^* \in \partial f(x)$,

(b) $\langle z, x^*\rangle - f(z)$ *atteint sa borne supérieure en* $z = x$,

(c) $f(x) + f^*(x^*) \leq \langle x, x^*\rangle$,

(d) $f(x) + f^*(x^*) = \langle x, x^*\rangle$.

Si de plus $f(x) = \overline{f}(x)$, *ces assertions sont encore équivalentes à :*

(a^*) $x \in \partial f^*(x^*)$,

(a^{**}) $x^* \in \partial \overline{f}(x)$,

(b^*) $\langle x, z^*\rangle - f^*(z^*)$ *atteint sa borne supérieure en* $z^* = x^*$.

L'assertion (b) signifie que $f^*(x^*) = \langle x, x^*\rangle - f(x)$, et l'assertion (b^*) signifie que $f(x) = \overline{f}(x) = f^{**}(x) = \langle x, x^*\rangle - f^*(x^*)$. Finalement, l'assertion (d) dit qu'en un point x donné, les vecteurs x^* pour lesquels l'inégalité de Young-Fenchel devient une égalité sont les sous-gradients de f en ce point x.

Nous pouvons donc écrire cette expression suivante équivalente pour le sous-différentiel de f en x :

$$\partial f(x) = \{x^* \in I\!R^n \mid f(x) + f^*(x^*) = \langle x, x^*\rangle\}. \tag{1.6}$$

En effet, soit un élément z^* appartenant au sous-différentiel défini en (1.6), nous avons donc que

$$\begin{aligned} \langle x, z^*\rangle &= f(x) + f^*(z^*) \\ \langle x, z^*\rangle &= f(x) + \sup_{t \in I\!R^n} \{\langle t, z^*\rangle - f(t)\} \end{aligned}$$

d'où

$$\langle x, z^*\rangle \geq f(x) + \langle t, z^*\rangle - f(t) \quad \forall t \in I\!R^n$$

Donc

$$f(t) \geq f(x) + \langle t - x, z^*\rangle \quad \forall t \in I\!R^n$$

i.e. que z^* appartient au sous-différentiel de la Définition 1.5.3.

Les trois assertions (a), (a^*) et (a^{**}) permettent de déduire l'important corollaire suivant :

Corollaire 1.5.6 *Si* $f : I\!R^n \longrightarrow \overline{I\!R}$ *est une fonction convexe, propre et fermée, alors* ∂f^* *est l'inverse de* ∂f *au sens des opérateurs multivoques. Autrement dit :*

$$x \in \partial f^*(x^*) \Leftrightarrow x^* \in \partial f(x)$$

ou encore :

$$\partial f^* = (\partial f)^{-1}$$

Preuve

Soit $z^* \in (\partial f)^{-1}(y^*)$, ce qui est équivalent à $y^* \in \partial f(z^*)$.
Donc par la définition (1.6) du sous-différentiel, nous pouvons écrire que

$$\langle y^*, z^* \rangle = f(z^*) + f^*(y^*)$$

Puisque f est fermée, nous avons que

$$\langle y^*, z^* \rangle = f^{**}(z^*) + f^*(y^*)$$

Et, donc par la même définition du sous-différentiel, nous avons la conclusion recherchée, à savoir, $z^* \in \partial f^*(y^*)$. □

Etant donné une fonction convexe $f : I\!R^n \longrightarrow I\!R \bigcup \{+\infty\}$, que nous supposons bornée inférieurement, nous nous proposons de résoudre le problème

$$\begin{array}{ll} \text{minimiser} & f(x) \\ x \in I\!R^n & \end{array} \tag{1.7}$$

Lorsque f est différentiable, nous savons qu'en calculant son gradient, nous pouvons en trouver un minimum global (qui est unique si f est strictement convexe) :

$$\overline{x} \text{ est un minimum global de } f \Leftrightarrow \nabla f(\overline{x}) = 0$$

et cela a lieu si et seulement si l'hyperplan affine de $I\!R^{n+1}$ qui supporte l'épigraphe de f au point $(\overline{x}, f(\overline{x}))$ est «horizontal».

En examinant la définition de sous-différentiel de f, si nous pouvons trouver un $\overline{x} \in I\!R^n$ tel que $0 \in \partial f(\overline{x})$, alors nous aurons

$$f(z) \geq f(\overline{x}) + \langle 0, z - \overline{x} \rangle \quad \forall z \in I\!R^n$$

i.e.

$$f(z) \geq f(\overline{x}) \quad \forall z \in I\!R^n$$

et nous aurons ainsi trouvé un minimum global de la fonction f, c'est-à-dire résolu le problème d'optimisation convexe (1.7).

Dans le cas où le problème comporte des contraintes convexes, c'est-à-dire lorsqu'il est du type

$$\begin{cases} \text{minimiser} & f(x) \\ \text{sous contrainte} & x \in X \end{cases}$$

où $f : I\!R^n \longrightarrow I\!R \bigcup \{+\infty\}$ est convexe et bornée inférieurement sur le convexe non vide $X \subseteq I\!R^n$, nous pouvons utiliser la fonction indicatrice de X, $I_X(\cdot)$ pour nous ramener à un problème sans contraintes[2] :

$$\begin{array}{ll} \text{minimiser} & f(x) + I_X(x) \\ x \in I\!R^n & \end{array}$$

2. Cette idée est utilisée par Rockafellar

En effet, lorsque $x \notin X$, $I_X(x) = +\infty$ et donc le minimum ne se trouve certainement pas hors de X. De plus, lorsque $x \in X$, $I_X(x) = 0$ et donc $f(x) + I_X(x) = f(x)$ de sorte que minimiser $f(\cdot)$ sur X est parfaitement équivalent à minimiser $f(\cdot) + I_X(\cdot)$ sur $I\!R^n$.

Pour trouver la solution à ce problème, comme $f(\cdot) + I_X(\cdot)$ est encore convexe (c'est la somme de deux fonctions convexes), nous avons vu qu'il suffisait de trouver $\overline{x}$ tel que

$$0 \in \partial[f(\cdot) + I_X(\cdot)]\ (\overline{x})$$

Nous allons voir plus loin que sous une certaine condition (que nous pourrons interpréter comme une qualification des contraintes), nous pouvons écrire

$$\partial\,[f(\cdot) + I_X(\cdot)]\ (\overline{x}) = \partial f(\overline{x}) + \partial I_X(\overline{x}) = \partial f(\overline{x}) + N_X(\overline{x})$$

où $N_X(\overline{x})$ est le cône normal à X en $\overline{x}$.

Il s'agira donc de trouver un sous-gradient de f en $\overline{x}$, $v \in \partial f(\overline{x})$ tel que

$$-v \in N_X(\overline{x})$$

et pour le cas différentiable, nous avons $-v = -\nabla f(\overline{x})$:

$$-\nabla f(\overline{x}) \in N_X(\overline{x})$$

Remarquons que si $X = I\!R^n$ (pas de contraintes) alors $I_X(x) = 0\ \forall x$ et $N_X(\overline{x}) = \{0\}\ \forall \overline{x}$, nous retrouvons ainsi le résultat du premier cas :

$$0 \in \partial f(\overline{x})$$

Par la définition de sous-gradient d'une fonction convexe $f : I\!R^n \longrightarrow \overline{I\!R}$, nous avons

$$\partial(\lambda f)(x) = \lambda\ \partial f(x) \quad \forall x \in I\!R^n\ \forall \lambda > 0$$

En effet, si $\lambda > 0$, la fonction λf est encore convexe et si $x^* \in \partial(\lambda f)(x)$, alors

$$(\lambda f)\,(z) \geq (\lambda f)\,(x) + \langle x^*, z - x\rangle \quad \forall z$$

i.e.

$$\lambda\ f(z) \geq \lambda\ f(x) + \langle x^*, z - x\rangle \quad \forall z$$

Soit, en divisant par λ :

$$f(z) \geq f(x) + \langle \frac{x^*}{\lambda}, z - x\rangle \quad \forall z$$

Autrement dit, $\dfrac{x^*}{\lambda} \in \partial f(x)$, donc $x^* \in \lambda\ \partial f(x)$. Inversement, si $x^* \in \lambda\ \partial f(x)$, alors $\frac{x^*}{\lambda} \in \partial f(x)$ et le raisonnement inverse montre que $x^* \in \partial(\lambda f)(x)$.

Lorsque nous considérons le sous-différentiel de la fonction $-f$ où f est une fonction convexe, en rapport avec ce que nous venons de voir, nous pouvons chercher le sous-différentiel équivalent ne faisant intervenir que f.
Soit $x^* \in \partial(-f)(x)$, alors par la définition (1.6), nous pouvons écrire

$$\langle x, x^* \rangle = -f(x) + (-f)^*(x^*)$$

Par la Proposition 1.3.2, nous savons que $(-f)^*(x^*) = -f^*(-x^*)$, donc

$$\langle x, x^* \rangle = -f(x) - f^*(-x^*)$$

i.e.

$$\langle x, -x^* \rangle = f(x) + f^*(-x^*)$$

ce qui équivaut par la définition (1.6) à $-x^* \in \partial f(x)$ ou $x^* \in -\partial f(x)$.

Pour ce qui est du sous-différentiel d'une somme de fonctions convexes, la formule

$$\partial(f_1 + \ldots + f_m)(x) = \partial f_1(x) + \ldots + \partial f_m(x)$$

ne sera applicable que si les domaines des fonctions f_i se recoupent *suffisamment*. Cette condition sera la qualification des contraintes. Plus précisément :

Théorème 1.5.7 *(Séparation d'une somme de sous-différentiels) Soient $f_1, \ldots, f_m : I\!R^n \longrightarrow \overline{I\!R}$ des fonctions convexes propres et soit $f \stackrel{\text{déf}}{=} f_1 + \ldots + f_m$. Alors nous avons toujours*

$$\partial f_1(x) + \ldots + \partial f_m(x) \subset \partial f(x) \quad \forall x$$

Si les ensembles convexes ri (dom f_i) ($i = 1, \ldots, m$) ont un point en commun, i.e. *si $\bigcap_{i=1}^m$ ri (dom f_i) $\neq \emptyset$, alors*

$$\partial f(x) = \partial f_1(x) + \ldots + \partial f_m(x) \quad \forall x$$

Preuve
Nous ne prouvons ici que la première partie du théorème, et seulement pour deux sous-différentiels, i.e. $\partial f_1(x) + \partial f_2(x) \subset \partial(f_1 + f_2)(x)$.
Ainsi, par la Définition 1.5.3, nous avons

$$y^* \in \partial f_1(x) \Leftrightarrow f_1(t) \geq f_1(x) + \langle t - x, y^* \rangle \quad \forall t \in I\!R^n$$

et

$$z^* \in \partial f_2(x) \Leftrightarrow f_2(t) \geq f_2(x) + \langle t - x, z^* \rangle \quad \forall t \in I\!R^n$$

Donc

$$y^* + z^* \in \partial f_1(x) + \partial f_2(x) \Rightarrow f_1(t) + f_2(t) \geq f_1(x) + \langle t - x, y^* \rangle + f_2(x) + \langle t - x, z^* \rangle \quad \forall t \in I\!R^n$$

où la dernière inéquation est équivalente à

$$(f_1 + f_2)(t) \geq (f_1 + f_2)(x) + \langle t - x, y^* + z^* \rangle \quad \forall t \in I\!R^n$$

i.e., toujours par la même définition, $y^* + z^* \in \partial(f_1 + f_2)(x)$. □

Pour une introduction plus détaillée des sous-gradients et des sous-différentiels, voir [17], ainsi que le cours de programmation non linéaire de seconde licence (prof. NGUYEN V.H.).

1.6 Opérateurs maximaux monotones

Dans la suite, nous allons beaucoup nous intéresser aux opérateurs maximaux monotones qui permettent entre autres de représenter des problèmes d'optimisation. Un opérateur maximal monotone est une correspondance multivoque (*i.e.* qui à un point associe un ensemble), qui « ressemble » à une fonction croissante, et qui ne peut pas être prolongée en un autre opérateur monotone sur un domaine strictement plus grand. Nous examinerons les propriétés de ces opérateurs et nous en verrons des exemples importants.

Nous allons travailler ici avec deux espaces de Hilbert réels X et Y généraux, mais pour les applications, la plupart du temps, nous aurons $X = Y = \mathbb{R}^n$, et le fait de passer du cadre général (X et Y) au cadre plus concret ($\mathbb{R}^n$), ne posera aucun problème.

Concept d'opérateur

Considérons deux espaces de Hilbert sur le corps commutatif des réels X et Y.

Définition 1.6.1 *Un opérateur T sur X et à valeurs dans Y est une application notée $T : X \rightrightarrows Y$ et est un sous-ensemble du produit direct $X \times Y$.*

Selon cette définition, il semble plutôt que nous ayons défini le *graphe* de l'opérateur T en question. En réalité, nous ne ferons aucune distinction entre l'opérateur lui-même et son graphe. Pour assimiler ces deux notions, nous poserons

$$\forall x \in X,\ T(x) \overset{\text{déf}}{=} \{y \in Y \mid (x,y) \in Gr(T)\} \tag{1.8}$$

Grâce à cette convention, nous obtenons l'équivalence suivante :

$$y \in T(x) \Leftrightarrow (x,y) \in Gr(T)$$

Le graphe et l'opérateur seront tous deux quelquefois notés T.

Il s'agit de bien se faire à l'idée qu'à un point x de X, l'opérateur T peut associer un *ensemble* de points de Y. C'est pourquoi nous écrivons $y \in T(x)$ et non pas $y = T(x)$. T n'est pas nécessairement univoque.

Définition 1.6.2 *Etant donné un opérateur $T : X \rightrightarrows Y$, son* inverse *est l'opérateur noté $T^{-1} : Y \rightrightarrows X$ et défini par*

$$(y,x) \in Gr(T^{-1}) \overset{\text{déf}}{\Leftrightarrow} (x,y) \in Gr(T)$$

ou encore

$$x \in T^{-1}(y) \overset{\text{déf}}{\Leftrightarrow} y \in T(x)$$

Par la définition, nous voyons immédiatement que l'inverse de l'opérateur T^{-1} est l'opérateur T de départ, *i.e.* $(T^{-1})^{-1} = T$. Contrairement à l'inverse d'une fonction, l'inverse d'un opérateur existe toujours.

Il se peut que pour un certain $x \in X$, l'ensemble image $T(x)$ soit vide. Alors, tout comme nous avons défini plus haut le domaine effectif d'une fonction convexe, définissons le *domaine* d'un opérateur :

Définition 1.6.3 *Etant donné un opérateur $T : X \rightrightarrows Y$, le* domaine *de T est l'ensemble*

$$D(T) \stackrel{\text{déf}}{=} \{x \in X \mid T(x) \neq \emptyset\} = \{x \in X \mid \exists\, y \in Y\ (x, y) \in T\}$$

et l' image *de T est l'ensemble*

$$im\, T \stackrel{\text{déf}}{=} \{y \in Y \mid \exists\, x \in X\ (x, y) \in T\} = D(T^{-1})$$

Un tel opérateur T pouvant associer une partie de Y à un point de X sera dit *multivoque*. Lorsque pour chaque $x \in X$, $T(x)$ est nécessairement un singleton $\{y\}$, l'opérateur T sera dit *univoque* et pourra être assimilé à une fonction. Dans ce cas, nous écrirons abusivement $y = T(x)$ plutôt que $\{y\} = T(x)$ ou que $y \in T(x)$.

Remarque 1.6.4 *Lorsque nous reprendrons le cas concret $X = Y = I\!R^n$, pour désigner un opérateur éventuellement multivoque et afin de ne pas le confondre avec une fonction sur $I\!R^n$, nous noterons $T : I\!R^n \rightrightarrows I\!R^n$ au lieu de $T : I\!R^n \longrightarrow I\!R^n$.*

Exemples

1. La correspondance $I : I\!R^n \rightrightarrows I\!R^n$ donnée par $x \rightsquigarrow \{x\}$, qui peut être assimilée à l'identité, est un opérateur comme défini plus haut. Son graphe est la diagonale de $I\!R^{2n}$: $\Delta_{I\!R^{2n}} = \{(x, x) \mid x \in I\!R^n\}$, son domaine est $I\!R^n$ tout entier, ainsi que son image. C'est un opérateur univoque.

2. Considérons une fonction convexe propre $f : I\!R^n \rightrightarrows \overline{I\!R}$. La correspondance $\partial f : I\!R^n \rightrightarrows I\!R^n$ est un opérateur comme défini plus haut. Il est univoque si et seulement si f est différentiable. Comme nous savons par le Théorème 1.5.4 que

$$x \notin \text{dom } f \Rightarrow \partial f(x) = \emptyset$$

et que

$$x \in \text{ri dom } f \Rightarrow \partial f(x) \neq \emptyset$$

le domaine de $\partial f(\cdot)$ est contenu dans *dom* f et contient *ri dom* f, i.e. *ri dom* $f \subset D(\partial f) \subset$ *dom* f.

3. Soit $C \subseteq I\!R^n$ une partie convexe. Le cône normal

$$N_C : I\!R^n \rightrightarrows I\!R^n$$

$$N_C(\overline{x}) = \begin{cases} \{v \in I\!R^n \mid \langle v, x - \overline{x}\rangle \leq 0 \ \forall x \in C\} & \text{si } \overline{x} \in C \\ \emptyset & \text{sinon} \end{cases}$$

est un opérateur comme défini plus haut, qui n'est univoque que si $C = I\!R^n$. Cet opérateur est de domaine C.

Opérations sur les opérateurs

Somme de deux opérateurs

Si $T_1, T_2 : X \rightrightarrows Y$ sont deux opérateurs, leur *somme* est définie par

$$(T_1 + T_2)(x) \stackrel{\text{déf}}{=} T_1(x) + T_2(x) \stackrel{\text{déf}}{=} \{t_1 + t_2 \mid t_1 \in T_1(x),\ t_2 \in T_2(x)\}$$

où cette dernière somme est la somme de deux *parties* de Y. Nous avons donc

$$T_1 + T_2 = \{(x, t_1 + t_2) \mid (x, t_1) \in T_1,\ (x, t_2) \in T_2\}$$

Le domaine de $T_1 + T_2$ est $D(T_1) \bigcap D(T_2)$.

Multiplication d'un opérateur par un scalaire

Si $T : X \rightrightarrows Y$ est un opérateur et λ un nombre réel, le produit λT est un opérateur $X \rightrightarrows Y$ défini par :

$$(\lambda T)(x) \stackrel{\text{déf}}{=} \lambda\, T(x) \stackrel{\text{déf}}{=} \{\lambda t \mid t \in T(x)\}$$

ou

$$\lambda T = \{(x, \lambda t) \mid (x, t) \in T\}$$

Le domaine de λT est le domaine de T.
Ce produit correspond donc à une «mise à l'échelle» de $im\ T$ en multipliant tous ses éléments par λ.

Somme d'un opérateur et d'un vecteur

Si $T : X \rightrightarrows Y$ est un opérateur et $s \in Y$ est un vecteur, leur somme $T + s = s + T$ est définie par

$$(T + s)(x) \stackrel{\text{déf}}{=} T(x) + s \stackrel{\text{déf}}{=} \{t + s \mid t \in T(x)\}$$

ou

$$T + s = \{(x, t + s) \mid (x, t) \in T\}$$

Cette somme correspond à une translation de $im\ T$ selon le vecteur s. Le domaine de $T + s$ est égal au domaine de T.

Composition avec une application linéaire

Si $T : X \rightrightarrows Y$ est un opérateur, si $M : Y \longrightarrow Z$ est une application linéaire, où Z est un espace de Hilbert, la composition $M \circ T$ est définie par :

$$(M \circ T)(x) \overset{\text{noté}}{=} M\,T(x) \overset{\text{déf}}{=} \bigcup_{y \in T(x)} My$$

De même, si $L : Z \longrightarrow X$ est une application linéaire, la composition $T \circ L$ est définie par $(T \circ L)(z) \overset{\text{noté}}{=} T(Lz)$.

Concept d'opérateur monotone

En analyse réelle, une fonction $f : I\!R \longrightarrow I\!R$ est dite *croissante* si et seulement si

$$\forall x, y \quad x \leq y \Rightarrow f(x) \leq f(y)$$

Cette implication peut encore se traduire par :

$$\forall x, y\ (x - y)(f(x) - f(y)) \geq 0 \tag{1.9}$$

ou encore par le fait que $(x - y)$ et $f(x) - f(y)$ ont toujours même signe. Nous allons généraliser cette idée aux opérateurs.

Nous nous plaçons ici dans le cas particulier où les deux espaces de Hilbert sont égaux : $Y = X$.
X étant un espace de Hilbert, il est muni d'un produit scalaire que nous noterons $\langle \cdot, \cdot \rangle$ et dont nous pouvons déduire une norme, notée $\| \cdot \|$.

Définition 1.6.5 *Un opérateur univoque $T : X \rightrightarrows X$ sera dit* monotone *si et seulement si*

$$\forall x, y \in X\ \langle x - y, T(x) - T(y) \rangle \geq 0 \tag{1.10}$$

Cette définition généralise la définition d'application croissante en ce sens que les deux vecteurs $x - y$ et $T(x) - T(y)$ ont « même signe ». Dans le cas où $X = I\!R$, nous voyons que (1.10) se ramène à (1.9). Cette définition s'étend directement aux opérateurs multivoques de la façon suivante :

Définition 1.6.6 *Un opérateur $T : X \rightrightarrows X$ sera dit* monotone *si et seulement si*

$$\forall x, x' \in X \quad \forall y \in T(x) \quad \forall y' \in T(x') \quad \langle x - x', y - y' \rangle \geq 0 \tag{1.11}$$

ce qui peut encore s'écrire

$$\forall (x, y), (x', y') \in T\ \langle x - x', y - y' \rangle \geq 0 \tag{1.12}$$

Dans le cas où l'opérateur T est univoque, (1.12) se ramène à (1.10) puisque $T(x) = \{u\}$ et $T(y) = \{v\}$. Remarquons qu'il faut ici comprendre « monotone » comme « croissant ». En analyse réelle, une fonction décroissante est aussi appelée monotone, mais elle ne vérifie pas (1.9); pour elle, l'inégalité doit être renversée.

Exemples

1. L'opérateur $I : I\!R^n \rightrightarrows I\!R^n$ donné par $x \rightsquigarrow \{x\}$ est monotone puisque[3]

$$\forall x, y \in I\!R^n \; \langle x - y, I(x) - I(y) \rangle = \langle x - y, x - y \rangle = \|x - y\|^2 \geq 0$$

2. Soit $f : I\!R^n \longrightarrow \overline{I\!R}$ une fonction convexe, propre. Considérons à nouveau l'opérateur sous-différentiel ∂f. Il est monotone car si $x, y \in I\!R^n$ et $u \in \partial f(x), v \in \partial f(y)$, nous avons

$$u \in \partial f(x) \Leftrightarrow f(z) \geq f(x) + \langle u, z - x \rangle \; \forall z$$
$$v \in \partial f(y) \Leftrightarrow f(w) \geq f(y) + \langle v, w - y \rangle \; \forall w$$

Prenons $z = y, w = x$ et additionnons ces inégalités :

$$f(y) + f(x) \geq f(x) + f(y) + \langle u, y - x \rangle + \langle v, x - y \rangle$$

i.e.

$$\langle x - y, u - v \rangle \geq 0$$

3. Le cône normal à une partie convexe C est aussi monotone car $\forall \overline{x}, \overline{y} \in C \; \forall u \in N_C(\overline{x}) \; \forall v \in N_C(\overline{y})$,

$$u \in N_C(\overline{x}) \Leftrightarrow \langle u, z - \overline{x} \rangle \leq 0 \; \forall z \in C$$
$$v \in N_C(\overline{y}) \Leftrightarrow \langle v, w - \overline{y} \rangle \leq 0 \; \forall w \in C$$

Prenons $z = \overline{y}, w = \overline{x}$ (qui sont dans C par hypothèse), et sommons :

$$\langle u, \overline{y} - \overline{x} \rangle + \langle v, \overline{x} - \overline{y} \rangle \leq 0$$

i.e.

$$\langle \overline{x} - \overline{y}, u - v \rangle \geq 0$$

4. Si nous reprenons l'opérateur I du premier point, nous voyons que $-I$ n'est pas monotone.

3. pour le produit scalaire, nous identifions $I(x)$, qui est $\{x\}$ et x lui-même.

5. Soit une fonction $f : I\!R \longrightarrow I\!R$ croissante, non-nécessairement continue. Définissons l'opérateur $\tilde{f} : I\!R \rightrightarrows I\!R$ par : $\tilde{f}(x) \stackrel{\text{déf}}{=} [f(x^-), f(x^+)]$. Cet opérateur est monotone. En effet, soient $x, y \in I\!R$ et $u \in \tilde{f}(x), v \in \tilde{f}(y)$;

$$u \in \tilde{f}(x) \Leftrightarrow -\infty < f(x^-) \leq u \leq f(x^+) < +\infty$$

$$v \in \tilde{f}(y) \Leftrightarrow -\infty < f(y^-) \leq v \leq f(y^+) < +\infty$$

Sans perte de généralité, nous pouvons supposer $x < y$. Alors

$$f(x^-) \leq f(x^+) \leq f(y^-) \leq f(y^+)$$

car f est croissante. Donc $u - v \leq 0$ et $x - y < 0$ d'où $(x-y)(u-v) \geq 0$.
Nous pouvons d'ailleurs montrer que tout opérateur monotone sur $I\!R$ est inclus dans un opérateur de ce type.

Propriétés des opérateurs monotones

Opérateur inverse

Si $T : X \longrightarrow X$ est un opérateur monotone, alors T^{-1} l'est aussi. Cela est simplement dû à la symétrie du produit scalaire dans (1.12). L'équivalence T monotone $\Leftrightarrow$ T^{-1} monotone rejoint bien l'intuition selon laquelle si $f : I\!R \longrightarrow I\!R$ est une fonction croissante et inversible (*i.e.* strictement croissante), alors f^{-1} est aussi croissante. Cela se voit graphiquement car le graphe de f^{-1} est le symétrique de celui de f par rapport à la diagonale principale.

Somme de deux opérateurs

Si $T_1, T_2 : X \longrightarrow X$ sont monotones, $T_1 + T_2 : X \longrightarrow X$ l'est aussi. En effet, si $x, y \in X$ et $u \in (T_1 + T_2)(x), v \in (T_1 + T_2)(y)$, nous avons $u = u_1 + u_2$ et $v = v_1 + v_2$ avec $u_1 \in T_1(x), u_2 \in T_2(x), v_1 \in T_1(y), v_2 \in T_2(y)$ et dès lors :

$$\begin{aligned} \langle x - y, u - v \rangle &= \langle x - y, (u_1 - v_1) + (u_2 - v_2) \rangle \\ &= \underbrace{\langle x - y, u_1 - v_1 \rangle}_{\geq 0} + \underbrace{\langle x - y, u_2 - v_2 \rangle}_{\geq 0} \geq 0 \end{aligned}$$

car T_1 et T_2 sont monotones.

Multiplication par un scalaire positif

Si $T : X \longrightarrow X$ est monotone et si $\lambda \geq 0$, alors λT l'est aussi. En effet, nous pouvons supposer que $\lambda > 0$ et si $x, y \in X$ $u \in (\lambda T)(x), v \in (\lambda T)(y)$, alors $\frac{u}{\lambda} \in T(x), \frac{v}{\lambda} \in T(y)$ et donc :

$$\langle x - y, u - v \rangle = \lambda \, \langle x - y, \frac{u}{\lambda} - \frac{v}{\lambda} \rangle \geq 0$$

car T est monotone et $\lambda > 0$.

Somme d'un opérateur et d'un vecteur

Si $T : X \longrightarrow X$ est monotone et si $s \in X$, alors $T + s$ l'est aussi. Pour le voir, prenons $x, y \in X$ et $u \in (T + s)(x), v \in (T + s)(y)$. Alors, par définition, nous avons $u - s \in T(x), v - s \in T(y)$. Donc :

$$\langle x - y, u - v\rangle = \langle x - y, (u - s) - (v - s)\rangle \geq 0$$

Définition 1.6.7 *Un opérateur T sur X est dit* fortement monotone *si*

$$\exists\ \alpha > 0,\ \forall x, y \in X \quad \langle T(y) - T(x), y - x\rangle \geq \alpha\|y - x\|^2$$

Définition 1.6.8 *Un opérateur T sur X est dit* non-expansif *si*

$$\|y' - y\| \leq \|x' - x\| \quad \forall (x, y), (x', y') \in Gr(T) \tag{1.13}$$

L'opérateur $(I + \lambda T)^{-1}$, noté J_λ, appelé *résolvante* de T, où I est l'opérateur identité, est non-expansive pour tout $\lambda > 0$.
Remarquons que $D(I + \lambda T) = D(T)$.

Le fait que J_λ soit non-expansive implique que pour tout y dans X, l'équation en x

$$y \in (I + \lambda T)(x) \tag{1.14}$$

admet *au plus* une solution.

A présent, nous allons considérer les opérateurs pour lesquels l'équation (1.14) possède *exactement* une solution.

Concept d'opérateur maximal monotone

En mathématiques classiques, la dernière hypothèse de base est liée à la question de savoir si nous pouvons toujours comparer deux ensembles. Cette hypothèse s'appelle *axiome du choix*. Elle ne nous sera pas utile sous sa forme de départ mais nous pouvons montrer qu'elle est équivalente au lemme de Zorn :

Lemme 1.6.9 (Lemme de Zorn) *Tout ensemble ordonné inductif non vide admet un élément maximal.*

Dans cet énoncé, « inductif » signifie que toute partie totalement ordonnée admet un majorant.

Notons $\mathcal{M}(X)$ l'ensemble des opérateurs monotones $T : X \rightrightarrows X$. Nous pouvons munir $\mathcal{M}(X)$ de l'ordre donné par l'inclusion des graphes. Cela signifie que

$$T_1 \leq T_2 \Leftrightarrow T_1 \subseteq T_2 \Leftrightarrow \forall x \in X\ T_1(x) \subseteq T_2(x)$$

Montrons que cet ensemble ordonné est inductif.
Si $\mathcal{P}$ en est une partie totalement ordonnée, alors nous avons par définition d'ordre total :

$$\forall T_1, T_2 \in \mathcal{P}\ T_1 \subseteq T_2 \text{ ou } T_2 \subseteq T_1$$

Si nous posons $\overline{T} \stackrel{\text{déf}}{=} \bigcup_{T \in \mathcal{P}} T$, alors $\overline{T}$ est un majorant de $\mathcal{P}$.

Comme $\mathcal{M}(X)$ est non vide, si nous acceptons l'axiome du choix, nous aurons aussi le lemme de Zorn, qui va permettre de définir les opérateurs maximaux.

Définition 1.6.10 *Un opérateur monotone $T : X \rightrightarrows X$ est dit* maximal *si c'est un élément maximal de $\mathcal{M}(X)$. Cela signifie que son graphe est maximal pour l'inclusion dans l'ensemble des graphes monotones, ou encore que :*

$$(x, u) \in X \times X \ \textit{et}\ \langle x - y, u - v \rangle \geq 0\ \forall (y, v) \in Gr(T) \Rightarrow (x, u) \in Gr(T)$$

Par cette définition, et par symétrie du produit scalaire, nous voyons que

$$T \text{ est maximal monotone} \Leftrightarrow T^{-1} \text{ est maximal monotone}$$

La proposition suivante est une caractérisation des opérateurs maximaux monotones.

Théorème 1.6.11 *Soit $T : X \rightrightarrows X$ un opérateur. Les assertions suivantes sont équivalentes :*

(i) T est maximal monotone,

(ii) T est monotone et im $(I + T) = X$ (Minty [12]),

(iii) Pour tout $\lambda > 0$, $J_\lambda = (I + \lambda T)^{-1}$ est un opérateur non-expansif défini sur X tout entier.

Preuve : Voir [4].

Il est facile de voir par la définition que si T est maximal monotone et si $\lambda > 0$, alors λT est encore maximal monotone. Dès lors, par le point *(ii)* du théorème, im $(I + \lambda T) = X$ et donc, pour tout $y \in X$, l'équation en x (1.14)

$$y \in (I + \lambda T)(x)$$

admet au moins une solution. Comme nous avons vu précédemment qu'elle en admettait au plus une, elle en admet une et une seule. Grâce à cette conclusion, le point *(iii)* du théorème se réécrit comme suit :

La résolvante $J_\lambda = (I + \lambda T)^{-1}$ d'un opérateur maximal monotone T sur X est univoque, non-expansive et partout définie.

Nous pouvons dire également que l'opérateur sous-différentiel d'une fonction convexe propre *et semi-continue inférieurement* n'est pas seulement monotone, il est maximal monotone.
Nous avons aussi que les opérateurs $T + I$ et $T - I$ sont maximaux monotones.

Définition 1.6.12 *Un opérateur T sur X est* borné au voisinage de x_0 *s'il existe un voisinage $\mathcal{U}$ de x_0 tel que $\bigcup_{x \in \mathcal{U}} Tx$ soit borné.*
Un opérateur T est localement borné *si T est borné au voisinage de tous les points de $D(T)$.*

Remarquons que si le produit d'un opérateur maximal monotone par un scalaire strictement positif est encore maximal monotone, il n'en va pas de même pour la somme de deux opérateurs maximaux monotones.

Nous avons vu dans le cadre des opérateurs sous-différentiels que si $f_1, f_2 : I\!R^n \longrightarrow \overline{I\!R}$ sont deux fonctions convexes, propres et que si ri dom $f_1 \bigcap$ ri dom $f_2 \neq \emptyset$, alors

$$\partial(f_1 + f_2) = \partial f_1 + \partial f_2$$

Or nous savons que si f_1 et f_2 sont convexes, propres et semi-continues inférieurement, alors $f \stackrel{\text{déf}}{=} f_1 + f_2$ est encore convexe, propre et semi-continue inférieurement.
Les opérateurs $T_1 \stackrel{\text{déf}}{=} \partial f_1$, $T_2 \stackrel{\text{déf}}{=} \partial f_2$ et $T \stackrel{\text{déf}}{=} \partial f$ sont donc maximaux monotones, et puisque, par le Théorème 1.5.4, nous avons que *ri dom* $g \subseteq D(\partial g)$ pour toute fonction g convexe, propre, ce qui entraîne *ri dom* $g \subseteq$ *ri* $D(\partial g)$, donc :

$$ri\ dom\ f_1 \bigcap ri\ dom f_2 \neq \emptyset \Rightarrow ri\ D(T_1) \bigcap ri\ D(T_2) \neq \emptyset$$

mais le membre de gauche de cette implication entraîne

$$\partial f = \partial f_1 + \partial f_2 = T_1 + T_2$$

et donc $T_1 + T_2$ est maximal monotone, puisque T l'est par hypothèse. Ceci justifie intuitivement le fait que la condition

$$ri\ D(T_1) \bigcap ri\ D(T_2) \neq \emptyset$$

soit *suffisante* pour que la somme $T_1 + T_2$ de deux opérateurs maximaux monotones soit un opérateur maximal monotone.

Un développement plus complet de la recherche de la monotonie maximale de la somme de deux opérateurs maximaux monotones est un des objets traité au chapitre suivant (voir Sous-section 2.2.3).

Exemples

1. L'opérateur identité I est maximal monotone. En effet, nous savons déjà qu'il est monotone, et comme im $(I + I)$ = im $2I = I\!R^n$, nous concluons par le point *(ii)* du Théorème 1.6.11 qu'il est maximal monotone.

2. La fonction $I_C(\cdot)$ indicatrice de la partie convexe non vide et *fermée* C est convexe, propre et semi-continue inférieurement (car C est fermée). Par la remarque précédente, nous avons donc que l'opérateur cône normal $N_C(\cdot) = \partial I_C(\cdot)$ est maximal monotone.

3. L'opérateur univoque (autre que l'identité) $T : I\!R^2 \longrightarrow I\!R^2$ défini par $T(x) = Ax$ où

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

   est maximal monotone.
   Cet opérateur est bien monotone car

$$\forall (x_1, x_2) \in I\!R^2 \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ -x_1 \end{pmatrix}$$

   L'image d'un vecteur lui est donc orthogonale. Nous en déduisons :

$$\forall x = (x_1, x_2), y = (y_1, y_2) \in I\!R^2 \ \langle x - y, Ax - Ay \rangle = \langle x - y, A(x - y) \rangle = 0 \geq 0$$

   Pour voir que T est aussi maximal, utilisons la définition. Si nous savons que $(x, u) \in I\!R^2 \times I\!R^2$ est tel que $\forall (y, Ay) \ \langle x - y, u - Ay \rangle \geq 0$, en écrivant cette relation avec toutes les composantes des vecteurs, nous avons

$$\forall (y_1, y_2) \ (x_1 - y_1)(u_1 - y_2) + (x_2 - y_2)(u_2 + y_1) \geq 0$$

   ou, en rassemblant les termes en y_1 et y_2 :

$$\forall (y_1, y_2) \ x_1 u_1 + x_2 u_2 \geq y_2(x_1 + u_2) - y_1(x_2 - u_1)$$

   Prenons $y_1 = 0$. Il faut montrer que $(u_1, u_2) = (x_2, -x_1)$. Supposons par l'absurde que $x_1 > -u_2$, alors, en faisant tendre y_2 vers $+\infty$, nous avons $\langle x, u \rangle \geq +\infty$ ce qui est absurde. Nous trouvons la même conclusion si nous supposons $x_1 < -u_2$ en faisant tendre y_2 vers $-\infty$. Nous avons donc $u_2 = -x_1$. Nous trouvons de la même façon $u_1 = x_2$. Par conséquent, nous avons obtenu $u = Ax$, ce qui montre que T est maximal monotone.

4. L'opérateur $f : [0, 1] \longrightarrow I\!R$ donné par $x \rightsquigarrow 2x$ est monotone, mais il n'est pas maximal monotone sur $I\!R$ car si nous prolongeons son graphe en étendant son domaine à $I\!R$ tout entier, nous obtenons un graphe monotone qui contient strictement celui de f. Ce dernier graphe *est* maximal sur $I\!R$.

Pour des rappels théoriques plus détaillés au sujet des opérateurs dans les espaces de Hilbert, voir [4].

Chapitre 2

Dualité générale

2.1 Dualité

2.1.1 Introduction

En théorie, la dualité sous ses différentes formes a participé longuement dans le développement de l'optimisation en général. La dualité nous aide à résoudre des problèmes de dimensions importantes sous leur forme primale, en passant à un problème de dimension moindre par le changement en son dual. Nous en présentons le principe général, et l'appliquons à un problème de la littérature bien connu : le problème de Fenchel-Rockafellar.

Récemment, Attouch et Théra [1] ont introduit et analysé un principe nouveau de dualité dit : Dualité Abstraite, que nous présenterons dans les sections qui suivent. Ce nouveau principe est très avantageux car il unifie un grand nombre de principes de dualité connus dans la littérature. Nous soulevons une remarque importante, dûe à l'ordre dans lequel nous considérons les opérateurs : une question quant à l'écriture du résultat après l'application du nouveau principe. Nous nous intéresserons ensuite aux opérateurs maximaux monotones et nous verrons en quoi ce résultat a amélioré la recherche de la somme d'opérateurs maximaux monotones.

2.1.2 Problèmes d'Optimisation

Soit le problème de minimisation suivant :

$$(P) \qquad \begin{array}{l} \inf f(x) \\ x \in X \end{array}$$

que nous appelons problème primal et nous notons (P).
Avec $X = I\!R^n$ et $f : I\!R^n \longrightarrow \overline{I\!R}$ est une fonction convexe, propre, semi-continue inférieurement.

Nous introduisons ensuite une fonction $F : X \times Y \longrightarrow \overline{I\!R}$ convexe, propre, appelée fonction de *perturbation* et qui est telle que $F(x,0) = f(x)\ \forall x \in X$. L'espace $Y = I\!R^m$ est appelé l'espace des perturbations.

Il est clair que le problème suivant

$$(P_{F,y}) \qquad \begin{array}{l} \inf F(x,y) \\ x \in X \end{array}$$

est équivalent à (P) pour $y = 0$, c'est-à-dire :

$$(P_{F,0}) \qquad \begin{array}{l} \inf F(x,0) \\ x \in X \end{array}$$

Un exemple de fonction de perturbation (ici, $X = Y = I\!R^n$)

$$F(x,y) = f(x+y)\ \ \forall x,y \in X$$

Pour obtenir le problème dual de $P_{F,y}$, nous utilisons la fonction conjuguée de F. Rappelons que pour une fonction convexe f, sa fonction conjuguée est définie comme suit (voir la Définition 1.3.1) :

$$f^*(x^*) = \sup_{x \in X}\{x^T x^* - f(x)\}$$

Donc, en transposant cette définition à la fonction de perturbation $F(x,y)$, nous obtenons

$$F^*(w^*,z^*) = \sup_{(x,y)\in X \times Y}\{(x,y)^T(w^*,z^*) - F(x,y)\} \tag{2.1}$$

où, ici, le produit scalaire sur $X \times Y$ noté $<.,.>_{X\times Y}$ est défini par

$$\langle(.,.),(.,.)\rangle_{X\times Y} = \langle .,.\rangle_X + \langle .,.\rangle_Y$$

i.e. $\forall (x,y),(w,z) \in X \times Y$, nous avons

$$\langle(x,y),(w,z)\rangle = (x,y)^T(w,z) = x^T w + y^T z = \langle x,w\rangle + \langle y,z\rangle$$

Pour $w^* = 0$, notre équation (2.1) devient

$$F^*(0,z^*) = \sup_{(x,y)\in X\times Y}\{y^T z^* - F(x,y)\}$$

L'inégalité de Young-Fenchel (1.4) appliquée à la fonction de perturbation F nous donne

$$\langle(x^*,y^*),(x,y)\rangle \leq F(x,y) + F^*(x^*,y^*)\ \ \forall (x,y),(x^*,y^*) \in X \times Y$$

dans laquelle nous posons $y = 0$ et $x^* = 0$, nous obtenons donc

$$0 \leq F(x,0) + F^*(0,y^*)\ \forall x \in X, \forall y^* \in Y$$

i.e.

$$-F^*(0,y^*) \leq F(x,0) \ \forall x \in X, \forall y^* \in Y$$
$$-F^*(0,y^*) \leq f(x) \ \forall x \in X, \forall y^* \in Y$$

Le théorème de dualité faible (voir Théorème 1.3.3) affirme que

$$\sup_{y^* \in Y} \{-F^*(0,y^*)\} \leq \inf_{x \in X} F(x,0)$$

La partie gauche de la dernière inégalité représente la borne inférieure de $(P_{F,0})$, et nous le notons par $(D_{F^*,0})$, le problème dual de $(P_{F,0})$:

$$(D_{F^*,0}) \qquad \begin{array}{l} \sup\{-F^*(0,y^*)\} \\ y^* \in Y \end{array}$$

En maximisant cette borne inférieure, nous trouvons une meilleure solution obtenue à partir du problème dual et donc qui est inférieure à la solution de notre problème primal $(P_{F,0})$.

Avec cette expression du dual du problème primal, nous pouvons aussi prouver que le dual du dual est le primal sous l'hypothèse supplémentaire que la fonction de perturbation est fermée.

En effet, considérons une fonction F convexe, semi-continue inférieurement et fermée $(F = \overline{F})$, ce qui signifie que $F^{**} = \overline{F} = F$ et le problème de minimisation perturbé :

$$(P_{F,0}) \qquad \begin{array}{l} \inf F(x,0) \\ x \in X \end{array}$$

dont le problème dual s'exprime comme suit :

$$(D_{F^*,0}) \qquad \begin{array}{l} \sup\{-F^*(0,y^*)\} \\ y^* \in Y \end{array}$$

Or, nous pouvons écrire ce problème en utilisant les propriétés des infimum et des supremum

$$\sup_{y^* \in Y} \{-F^*(0,y^*)\} = -\inf_{y^* \in Y} F(0,y^*)$$

c'est-à-dire, par analogie

$$(-P_{F^*,0}) \qquad \begin{array}{l} -\inf F^*(0,y^*) \\ y^* \in Y \end{array}$$

Dualisons maintenant ce dernier problème, nous obtenons

$$(-D_{F^{**},0}) \qquad \begin{array}{l} -\sup\{-F^{**}(x^{**},0)\} \\ x^{**} \in X \end{array}$$

Nous pouvons également réécrire cette expression sous une autre forme en utilisant les hypothèses

$$\begin{aligned} -\sup_{x^{**}\in X}\{-F^{**}(x^{**},0)\} &= \inf_{x^{**}\in X} F^{**}(x^{**},0) \\ &= \inf_{x^{**}\in X} F(x^{**},0) \\ &= \inf_{x\in X} F(x,0) \end{aligned}$$

Nous sommes bien revenus au problème perturbé $(P_{F,0})$.

2.1.3 Problème de Fenchel-Rockafellar

Prenons maintenant comme fonction objectif f, la fonction définie par :

$$f(x) = g(x) + h(Ax)$$

où $g : X = I\!R^n \longrightarrow I\!R \cup \{+\infty\}$ et $h : Y = I\!R^m \longrightarrow I\!R \cup \{+\infty\}$ sont des fonctions convexes, semi-continues inférieurement et $A : X \longrightarrow Y$ est une application linéaire.
Le problème (P) devient donc :

$$(P1) \qquad \begin{array}{l} \inf\{g(x) + h(Ax)\} \\ x \in X \end{array}$$

Ce qui revient en réalité à considérer le problème soulevé par Fenchel et Rockafellar (voir le cours de programmation non linéaire de seconde licence (prof. NGUYEN V.H.) et [2]).

Lemme 2.1.1 *Soient $f : X \longrightarrow I\!R\cup\{+\infty\}$ et $g : Y \longrightarrow I\!R\cup\{+\infty\}$ deux fonctions convexes, semi-continues inférieurement et A une application linéaire de Y vers X. Alors,*

- *$f + g$ est convexe, semi-continue inférieurement,*
- *$f \circ A$ est convexe, semi-continue inférieurement.*

Ce lemme a été déduit d'un ensemble de propriétés des fonctions convexes, semi-continues inférieurement présentées dans [2].

Donc, la fonction $f = g + (h \circ A)$ est convexe et semi-continue inférieurement.

Calculons le problème dual correspondant à $(P1)$, pour cela, définissons une fonction de perturbation $F(x, y)$ de la manière suivante, i.e. ici nous utilisons une perturbation horizontale qui se définit comme ceci :

$$F(x, y) = g(x) + h(Ax + y)$$

Le problème $(P1)$ perturbé s'écrit donc :

$$(P1_{F,y}) \qquad \begin{array}{l} \inf F(x, y) \\ x \in X \end{array}$$

Nous aurions pu aussi perturber la fonction objectif horizontalement suivant le premier terme i.e.

$$F(x,y) = g(x-y) + h(Ax)$$

où $y \in X$, ou suivant les deux termes i.e.

$$F(x,y) = g(x-y_1) + h(Ax-y_2)$$

où $(y_1, y_2) \in X \times Y$.

Une perturbation verticale i.e., par exemple,

$$F(x,y) = g(x) + h(Ax) + y$$

où $y \in I\!R \cup \{+\infty\}$, n'est pas envisagée car l'ajout d'une constante ne change pas la recherche de l'infimum.

Cette fonction F est également convexe, semi-continue inférieurement, et pour $y = 0$, nous obtenons

$$F(x,0) = g(x) + h(Ax)$$

qui coïncide avec l'objectif de $(P1)$.
Le problème perturbé est donc défini comme suit :

$$(P1_{F,y}) \qquad \begin{array}{l} \inf\{g(x) + h(Ax+y)\} \\ x \in X \end{array}$$

Par substitution par rapport au problème (P), le problème dual de $(P1_{F,0})$ est équivalent à :

$$(D1_{F^*,0}) \qquad \begin{array}{l} \sup\{-F^*(0,y^*)\} \\ y \in Y \end{array}$$

Calculons la fonction conjuguée de F :

$$\begin{aligned} F^*(w^*,z^*) &= \sup_{(x,y)\in X\times Y} \{(x,y)^T(w^*,z^*) - F(x,y)\} \\ &= \sup_{(x,y)\in X\times Y} \{x^T w^* + y^T z^* - F(x,y)\} \\ &= \sup_{(x,y)\in X\times Y} \{x^T w^* + y^T z^* - g(x) - h(Ax+y)\} \end{aligned}$$

où $x, w^* \in X$ et $y, z^* \in Y$.
La fonction g étant définie seulement sur X, et le produit scalaire $x^T w^*$ ne dépendant que de x et de w^*, nous pouvons séparer comme suit :

$$F^*(w^*,z^*) = \sup_{x\in X}\{x^T w^* - g(x) + \sup_{y\in Y}\{y^T z^* - h(Ax+y)\}\}$$

$$\sup_{y\in Y}\{y^T z^* - h(Ax+y)\} = \sup_{y\in Y}\{(Ax+y)^T z^* - h(Ax+y) - (Ax)^T z^*\}$$

En posant $u = Ax + y$ ($u \in Y$ puisque $A : X \longrightarrow Y$ est une application linéaire), nous avons

$$\begin{aligned}
\sup_{y\in Y}\{(Ax+y)^T z^* - h(Ax+y) - (Ax)^T z^*\} &= \sup_{u\in Y}\{u^T z^* - h(u) - (Ax)^T z^*\} \\
&= \sup_{u\in Y}\{u^T z^* - h(u)\} - (Ax)^T z^* \\
&= h^*(z^*) - (Ax)^T z^*
\end{aligned}$$

La fonction conjuguée de h existe puisque nous avons défini la fonction h comme une fonction convexe et semi-continue inférieurement.

Reprenons le calcul de $F^*(w^*, z^*)$:

$$\begin{aligned}
F^*(w^*, z^*) &= \sup_{x\in X}\{x^T w^* - g(x) + h^*(z^*) - (Ax)^T z^*\} \\
&= \sup_{x\in X}\{x^T w^* - g(x) - (Ax)^T z^*\} + h^*(z^*)
\end{aligned}$$

Puisque la fonction h^* ne dépend que de l'argument z^*, la séparation peut prendre effet. De plus, si nous écrivons le produit scalaire comme $(Ax)^T z^* = x^T A^T z^*$, nous obtenons :

$$\begin{aligned}
F^*(w^*, z^*) &= \sup_{x\in X}\{(x^T w^* - g(x) - x^T A^T z^*\} + h^*(z^*) \\
&= \sup_{x\in X}\{(x^T (w^* - A^T z^*) - g(x)\} + h^*(z^*) \\
&= g^*(w^* - A^T z^*) + h^*(z^*)
\end{aligned}$$

où nous obtenons la dernière égalité par la définition de la fonction conjuguée de g.
Rappelons que la fonction conjuguée d'une fonction convexe semi-continue inférieurement est concave et semi-continue supérieurement.
Suivant les mêmes étapes que pour la recherche du problème dual de (P), i.e. poser, dans l'expression de $F^*(w^*, z^*)$, $w^* = 0$ et $z^* = y^*$, le problème dual de $(P1_{F,y})$ est donc :

$$(D1_{F^*,0}) \qquad \begin{array}{l} \sup\{-F^*(0, y^*)\} \\ y^* \in Y \end{array}$$

i.e.

$$(D1) \qquad \begin{array}{l} \sup\{-g^*(-A^T y^*) - h^*(y^*)\} \\ y^* \in Y \end{array}$$

Voici l'expression du problème dual du problème (P1), après le calcul de la conjuguée de la fonction de perturbation. Nous exprimons plus loin, dans la section suivante, un moyen plus rapide et moins calculatoire, autre que la dualité de Fenchel, pour obtenir son dual.

2.2 La dualité pour une somme d'opérateurs

Nous venons de traiter un problème de dualité impliquant des fonctions convexes et semi-continues inférieurement. Dans cette section, nous abordons un problème faisant intervenir des opérateurs (éventuellement maximaux monotones). Un résultat général de dualité abstraite est proposé pour des problèmes exprimés avec une somme d'opérateurs (éventuellement multivoques). Ce principe unifie une classe plus générale de principes de dualité connus. Cette somme est habituellement rencontrée en économie, en mécanique et en physique. A l'aide de ce principe de dualité abstraite, de nombreuses questions théoriques peuvent être soulevées en général, et en particulier pour le cas d'opérateurs maximaux monotones. Nous retraitons le problème soulevé par Rockafellar et Fenchel à l'aide des sous-différentiels, avec le nouveau principe général de dualité abstraite et nous en comparons les solutions avec les résultats précédents. Nous reprendrons la vision de Attouch-Riahi-Théra (introduite dans [1]) pour s'assurer de la maximalité de la somme de deux opérateurs maximaux monotnes. Enfin, nous présentons quelques résultats dûs à l'application de ce nouveau principe.

2.2.1 Le principe général de dualité abstraite

Définition 2.2.1 *(Principe de dualité abstraite : Attouch et Théra [1]) Soient X et Y deux espaces linéaires, et $A, B : X \rightrightarrows Y$ deux opérateurs (éventuellement multivoques).*

(a) Les deux problèmes variationnels :

$$(PV1) \qquad \alpha \in Ax + Bx$$

et

$$(PV2) \qquad 0 \in A^{-1}y - B^{-1}(\alpha - y)$$

sont équivalents dans ce sens :
Si x est une solution de (PV1), alors il existe $y \in Ax$ qui est solution de (PV2).
Si y est une solution de (PV2), alors il existe $x \in A^{-1}y$ qui est solution de (PV1).

(b) Le problème variationnel (PV2) est équivalent à

$$(PV3) \qquad \alpha \in y + BA^{-1}y$$

Remarque 2.2.2 *Le signe d'appartenance est préférable au signe d'égalité dans les inclusions (PV1), (PV2) et (PV3) et nous pouvons l'expliquer par le fait que les opérateurs A et B ne sont pas nécessairement univoques.*

Explication
(PV1) est équivalent à $\alpha - y \in Bx$, si nous prenons un y tel que $y \in Ax \neq \emptyset$. Cela revient à dire que x appartient à $A^{-1}y$ et à $B^{-1}(\alpha - y)$, càd $-x \in -B^{-1}(\alpha - y)$. Par addition, nous

avons le problème (PV2).
Reprenons ce dernier problème et écrivons le sous la forme

$$B^{-1}(\alpha - y) \in A^{-1}y$$

c'est-à-dire,

$$\alpha - y \in BA^{-1}y$$

ou encore

$$\alpha \in y + BA^{-1}y$$

en isolant α à gauche, c'est-à-dire le problème (PV3) $\diamondsuit$

Si nous prenons $\alpha = 0$ dans la définition 2.2.1, nous obtenons le résultat suivant :

Corollaire 2.2.3 *Les deux problèmes variationnels :*

$$(PV4) \qquad 0 \in Ax + Bx$$

$$(PV5) \qquad 0 \in A^{-1}y - B^{-1}(-y)$$

sont équivalents.
Le problème variationnel (PV5) est équivalent à

$$(PV6) \qquad 0 \in y + BA^{-1}y$$

Vérification : Le dual du dual est le primal
Exprimons que la double application de ce nouveau principe nous restitue le problème de départ (PV1).

Soit le problème variationnel

$$(PV1) \qquad \alpha \in Ax + Bx$$

équivalent par le principe de dualité abstraite à

$$(PV2) \qquad 0 \in A^{-1}y - B^{-1}(\alpha - y)$$

Simplifions au maximum ce problème, nous avons

$$\begin{aligned} B^{-1}(\alpha - y) \in A^{-1}y &\iff \alpha - y \in BA^{-1}y \\ &\iff \alpha \in y + BA^{-1}y \end{aligned}$$

Posons $C = I$ et $D = BA^{-1}$.
Le problème se réécrit donc comme suit

$$\alpha \in Cy + Dy$$

Nous pouvons y appliquer le principe de dualité abstraite puisque le problème s'exprime sous la même forme que (PV1). Nous avons donc

$$0 \in C^{-1}z - D^{-1}(\alpha - z)$$

Lorsque nous remplaçons C et D par leurs expressions, nous obtenons

$$0 \in (I^{-1})^{-1}z - (BA^{-1})^{-1}(\alpha - z)$$

i.e.

$$0 \in z - (BA^{-1})^{-1}(\alpha - z)$$

Simplifions de nouveau

$$\begin{aligned} z \in (BA^{-1})^{-1}(\alpha - z) &\iff \alpha - z \in BA^{-1}z \\ &\iff \alpha \in z + BA^{-1}z \end{aligned}$$

Si nous posons $z = Ax$, nous avons

$$\alpha \in Ax + Bx$$

nous obtenons donc le problème (PV1).
Ce qui montre que le dual au sens abstrait du dual (PV2) est le problème primal.

Remarque 2.2.4 *Rôle symétrique*
A et B ont un rôle symétrique.
En posant $z = -y$ dans le problème variationnel (PV5), nous obtenons

$$0 \in B^{-1}z - A^{-1}(-z) \tag{2.2}$$

qui est bien le problème (PV5), mais avec les rôles des opérateurs A et B inversés.
De la même façon, si nous prenons $z = -y$ dans (PV6), nous avons un problème équivalent à

$$0 \in -z + BA^{-1}(-z) \tag{2.3}$$

Donc, nous avons $z \in BA^{-1}(-z)$, qui est équivalent à $B^{-1}(z) \in A^{-1}(-z)$. Nous pouvons ensuite obtenir cette inclusion,

$$0 \in B^{-1}(z) - A^{-1}(-z)$$

L'expression suivante est $A^{-1}(-z) \in B^{-1}(z)$ qui nous donne $-z \in AB^{-1}(z)$.
Nous arrivons enfin à

$$0 \in z + AB^{-1}z \tag{2.4}$$

qui est également un problème exprimé comme (PV6), mais encore une fois avec les rôles de A et B inversés.

Remarque 2.2.5 *Somme de plusieurs opérateurs*
Nous pouvons également considérer un problème variationnel impliquant la somme de trois

ou de plusieurs opérateurs.
Considérons un exemple à trois opérateurs (le mécanisme étant identique au-delà).
Prenons $A, B, C : X \rightrightarrows Y$ *dans l'expression du problème variationnel suivant :*

$$0 \in Ax + Bx + Cx \tag{2.5}$$

Soit $\alpha \in Ax$ *et* $\beta \in Bx$*, i.e.* $x \in A^{-1}\alpha$ *et* $x \in B^{-1}\beta$*. Le problème (2.5) devient*

$$0 \in \alpha + \beta + Cx \tag{2.6}$$

Donc, nous avons que $(-\alpha - \beta) \in Cx$ *i.e.* $x \in C^{-1}(-\alpha - \beta)$*. En rassemblant toutes les informations précédentes, le problème (2.6) est équivalent aux deux inclusions suivantes :*

$$0 \in A^{-1}\alpha - C^{-1}(-\alpha - \beta)$$

et

$$0 \in B^{-1}\beta - C^{-1}(-\alpha - \beta)$$

qui sont équivalentes au problème (2.5) comme dans le principe général de dualité abstraite. Nous pouvons donc appliquer le principe à plusieurs opérateurs en utilisant le même mécanisme que pour trois opérateurs.
Nous pouvons maintenant chercher le zéro de plus de deux opérateurs en introduisant des éléments intermédiaires pour ne considérer que deux opérateurs le temps d'appliquer le principe de dualité abstraite.

Qu'y a-t-il de nouveau avec ce principe de dualité abstraite?

Un premier avantage visible directement de ce principe est que le long calcul de la fonction conjuguée de Fenchel n'est plus nécessaire pour obtenir le problème dual du problème primal. Ce changement est évident lorsque nous regardons l'expression du problème dual, qui n'utilise que les fonctions inverses des opérateurs.

Mais cela nous amème un autre problème, à savoir le calcul des inverses de fonctions convexes semi-continues inférieurement. Si en théorie, le principe paraît plus simple, car le calcul de la fonction conjuguée n'est pas aisé, en pratique, nous avons quand même des recherches à effectuer. Avant l'apparition de ce principe, nous devions d'abord choisir une fonction de perturbation adéquate et ensuite évaluer sa fonction conjuguée. Maintenant, il ne nous reste plus qu'à trouver la fonction inverse de la fonction objectif, le problème sera d'ordre technique dans certains cas.

Comme nous l'avons dit précédemment, ce nouveau principe unifie un grand nombre de principes de dualité bien connus. En appliquant le principe de dualité abstraite à un problème variationnel connu, nous obtenons un problème dual identique au problème dual obtenu par le principe habituellement appliqué à ce problème. Nous le vérifions pour le principe de dualité de Fenchel.

Nous avons remarqué une différence d'écriture lors de l'application de ce nouveau principe suivant le sens dans lequel nous considérons la somme du problème variationnel. Cette observation est l'objet de la section suivante. Nous illustrons cette remarque sur l'exemple de la dualité de Fenchel.

2.2.2 Observation du résultat

Nous avons souligné à la fin de la section précédente que le nouveau principe général de dualité abstraite amenait une différence d'écriture du résultat lors de son application. Comme nous le voyons dans l'expression du problème variationnel (PV1), il est écrit sous la forme d'une somme, et suivant le terme par lequel nous commençons pour appliquer le principe, nous n'obtenons pas la même expression du problème dual.

En effet, la somme des deux opérateurs peut s'écrire de deux façons différentes :

$$A + B = B + A$$

et donc en appliquant le principe de dualité à l'une ou à l'autre somme, nous n'obtenons pas la même écriture du résultat.

Soient les deux problèmes variationnels primaux équivalents suivants

$$0 \in Ax + Bx \Longleftrightarrow 0 \in Bx + Ax \tag{2.7}$$

auxquels nous appliquons le principe de dualité abstraite.
Nous obtenons donc ces deux inclusions :

$$0 \in A^{-1}y - B^{-1}(-y) \tag{2.8}$$

pour la partie gauche de l'équivalence (2.7), et

$$0 \in B^{-1}y - A^{-1}(-y) \tag{2.9}$$

pour la partie droite de l'équivalence (2.7). Ces deux dernières inclusions (2.8) et (2.9) représentent le même problème variationnel dual mais avec une écriture différente. La présence du signe négatif dans l'argument des opérateurs inverses de A et B fait toute la différence, mais les deux inclusions expriment le même problème.
En effet, si nous posons $z = -y$ dans l'inclusion (2.9), nous obtenons

$$0 \in B^{-1}(-z) - A^{-1}z$$

En le multipliant par (-1), il est équivalent à

$$0 \in A^{-1}z - B^{-1}(-z) \tag{2.10}$$

Lorsque nous regardons les inclusions (2.9) et (2.10), elles sont identiques. Le cas où les opérateurs A et B sont identiques est évidemment trivial.

Remarquons que si nous réappliquons le principe aux deux problèmes variationnels obtenus, nous voyons réapparaître le problème (PV1), ce qui est évident puisque nous avons prouvé précédemment que le dual du dual est le primal.

Reprenons maintenant le problème soulevé par Fenchel et Rockafellar et appliquons-y le principe de dualité abstraite.
Nous considérons ce problème à l'aide des inclusions de sous-différentiels pour pouvoir le lier plus facilement avec notre nouveau principe, car ce dernier comporte des inclusions d'opérateurs et nous avons vu à la Section 1.6, que les sous-différentiels sont des opérateurs maximaux monotones.

Pour confirmer la remarque précédente au sujet de la différence d'écriture du dual, nous cherchons le dual de ce problème pour les deux écritures de la somme des opérateurs, et nous obtenons ainsi deux problèmes duaux du problème de Fenchel et Rockafellar exprimés sous deux formes différentes.
Après avoir réécris les inclusions, nous les comparons au problème dual (D1) obtenu dans la section précédente. Une des deux inclusions doit être écrite strictement identique au problème (D1), l'autre doit s'en approcher mais les signes dans les arguments des opérateurs sont inversés.

Prenons le problème

$$(P1) \qquad \begin{array}{l} \inf\{g(x) + h(Ax)\} \\ x \in X \end{array}$$

Son problème dual, obtenu à la section précédente, est

$$(D1) \qquad \begin{array}{l} \sup\{-g^*(-A^Ty^*) - h^*(y^*)\} \\ y^* \in Y \end{array}$$

Le problème (P1) est équivalent par passage aux inclusions de sous-différentiels à

$$0 \in \partial(g + h \circ A)(x)$$

Supposons que

$$ri\ dom\ g \cap ri\ dom\ (h \circ A) \neq \emptyset$$

alors, par le Théorème 1.5.7, nous pouvons séparer le sous-différentiel comme suit

$$0 \in \partial g(x) + \partial(h \circ A)(x) \tag{2.11}$$

De même, le problème (D1) est équivalent à

$$0 \in \partial(-g^* \circ (-A^T) - h^*)(y^*)$$

Lorsque

$$ri\ dom\ (-g^* \circ (-A^T)) \cap ri\ dom\ (-h^*) \neq \emptyset$$

alors

$$0 \in \partial(-g^* \circ (-A^T))(y^*) + \partial(-h^*)(y^*) \tag{2.12}$$

Prenons le problème $(P1)$ sous la forme (2.11), et appliquons le principe de dualité abstraite, nous obtenons le problème équivalent suivant

$$0 \in (\partial g)^{-1}(y^*) - (\partial(h \circ A))^{-1}(-y^*) \tag{2.13}$$

Reprenons maintenant le problème (2.11), mais avec la somme des opérateurs inversée, i.e.

$$0 \in \partial(h \circ A)(x) + \partial g(x)$$

et appliquons-y le principe de dualité abstraite, nous obtenons

$$0 \in (\partial(h \circ A))^{-1}(y^*) - (\partial g)^{-1}(-y^*) \tag{2.14}$$

Comparons les problèmes (2.13) et (2.14). Ces deux inclusions sont écrites différemment : le signe négatif dans le sous-gradient ne se situe pas dans le même opérateur inverse.
De la même façon que pour le problème général, posons $z^* = -y^*$ dans (2.14), nous obtenons

$$0 \in (\partial(h \circ A))^{-1}(-z^*) - (\partial g)^{-1}(z^*)$$

et en multipliant par (-1), nous avons le même problème qu'à l'inclusion (2.13).

Utilisons les propriétés des sous-différentiels (Section 1.5) et des fonctions conjuguées (Section 1.3) pour transformer ces inclusions afin que leurs expressions soient plus proches du problème (2.12) et plus facilement comparables.

Considérons d'abord le problème (2.13).
Nous savons que l'inverse du sous-différentiel est équivalent au sous-différentiel de la fonction conjuguée (voir Corollaire 1.5.6), i.e.

$$(\partial f)^{-1} = \partial f^*$$

Donc le problème (2.13) devient

$$0 \in \partial g^*(y^*) - \partial(h \circ A)^*(-y^*) \tag{2.15}$$

Nous nous intéressons au deuxième terme de cette inclusion

$$\begin{aligned}
(h \circ A)^*(-y^*) &= \sup_{x \in X}\{\langle -y^*, x\rangle - (h \circ A)(x)\} \\
&= \sup_{x \in X}\{\langle -y^*, x\rangle - h(Ax)\} \\
&= \sup_{x \in X}\{\langle A^T(A^T)^{-1}(-y^*), x\rangle - h(Ax)\} \\
&= \sup_{z \in Y}\{\langle (A^T)^{-1}(-y^*), z\rangle - h(z)\} \\
&= h^*\left[(A^T)^{-1}(-y^*)\right]
\end{aligned}$$

obtenu par la définition des sous-différentiels et du produit scalaire.

Cherchons maintenant le sous-différentiel correspondant

$$\begin{aligned}
z^* \in \partial(h \circ A)^*(-y^*) &\Leftrightarrow \langle z^*, -y^* \rangle = (h \circ A)^*(-y^*) + (h \circ A)^{**}(z^*) \\
&\Leftrightarrow \langle z^*, -y^* \rangle = h^*((A^T)^{-1}(-y^*)) + (h \circ A)(z^*) \\
&\Leftrightarrow \langle z^*, -y^* \rangle = h^*((A^T)^{-1}(-y^*)) + h(Az^*) \\
&\Leftrightarrow \langle z^*, -y^* \rangle = h^*((A^T)^{-1}(-y^*)) + h^{**}(Az^*) \\
&\Leftrightarrow \langle Az^*, (A^T)^{-1}(-y^*) \rangle = h^*((A^T)^{-1}(-y^*)) + h^{**}(Az^*) \\
&\Leftrightarrow Az^* \in \partial h^*((A^T)^{-1}(-y^*)) \\
&\Leftrightarrow z^* \in A^{-1}\partial h^*((A^T)^{-1}(-y^*))
\end{aligned}$$

Ces équivalences sont correctes par la définition du sous-différentiel et les propriétés des fonctions convexes, propres, strictes.
En remplaçant ce que nous venons d'obtenir dans l'inclusion (2.15), nous avons

$$0 \in \partial g^*(y^*) - A^{-1}\partial h^*((A^T)^{-1}(-y^*)) \tag{2.16}$$

Nous pouvons utiliser les mêmes démarches pour transformer l'inclusion (2.14), nous aboutissons à

$$0 \in A^{-1}\partial h^*((A^T)^{-1}y^*) - \partial g^*(-y^*) \tag{2.17}$$

Ces deux dernières inclusions (2.16) et (2.17) ne sont pas écrites de façon identique, le signe négatif dans l'argument du sous-différentiel ne s'applique pas au même opérateur. Le résultat de l'application du principe général de dualité abstraite donne la même expression du dual.

Reprenons le problème (2.12), et utilisons les mêmes démarches que pour les inclusions précédentes, nous obtenons

$$\begin{aligned}
0 \in \partial(-g^* \circ (-A^T))(y^*) + \partial(-h^*)(y^*) &\Leftrightarrow 0 \in -A\partial(-g^*)(-A^T y^*) - \partial h^*(y^*) \\
&\Leftrightarrow 0 \in -A\partial g^*(-A^T y^*) + \partial h^*(y^*) \\
&\Leftrightarrow 0 \in -\partial g^*(-A^T y^*) + A^{-1}\partial h^*(y^*)
\end{aligned}$$

Si nous posons $z^* = A^T y^*$ dans la dernière inclusion, i.e. $y^* = (A^T)^{-1}z^*$, nous obtenons

$$0 \in -\partial g^*(-z^*) + A^{-1}\partial h^*((A^T)^{-1}z^*) \tag{2.18}$$

qui est identique à (2.17).
De même, si nous posons $z^* = -A^T y^*$, i.e. $y^* = (A^T)^{-1}(-z^*)$, nous obtenons

$$0 \in -\partial g^*(z^*) + A^{-1}\partial h^*((A^T)^{-1}(-z^*)) \tag{2.19}$$

qui est identique à (2.16) au signe près.

Nous avons obtenu le problème dual du problème $(P1)$, de deux manières différentes : la première par la méthode habituelle en calculant la fonction conjuguée de la fonction de perturbation et la deuxième par le nouveau principe de dualité abstraite.

Nous avons donc montré que le nouveau principe est fiable et que nous pouvons dorénavant l'appliquer pour chercher le problème dual de tout problème connu posé sous la forme d'une inclusion de somme de sous-différentiels ou faisant intervenir des opérateurs dans l'inclusion de départ comme dans (PV1).

Il permet donc d'unifier beaucoup de principes de dualité variationnels bien connus, nous n'avons développé ici que la dualité convexe de Fenchel-Moreau-Rockafellar.
Pour la présentation d'autres principes tels que le principe de Clarke-Ekeland ou la dualité de Singer-Toland, reportez-vous à l'article de Attouch et Théra [1], ces applications supplémentaires demandant une maitrise plus importante de connaissances telles que les systèmes hamiltoniens, les équations d'onde ou les équilibres de plasma.

De plus, ce nouveau principe permet de trouver le problème dual sans passer par le calcul assez long de la fonction conjuguée.

2.2.3 Au sujet de la somme d'opérateurs

Le principe général de dualité abstraite nous permet de démontrer un nouveau théorème concernant la somme de deux opérateurs maximaux monotones et donc de savoir, en général, sous quelles conditions la somme de deux opérateurs maximaux monotones est un opérateur maximal monotone.

Les opérateurs maximaux monotones généraux

Jusqu'ici, nous ne connaissons que quelques théorèmes concernant la somme de deux opérateurs maximaux monotones.
Ces théorèmes font intervenir deux opérateurs monotones mais qui ne sont pas nécessairement tous les deux maximaux. Brézis [4] nous donne quelques pistes pour obtenir un opérateur-somme maximal monotone mais, soit il n'obtient pas la maximalité de l'opérateur-somme, soit la condition sur un des deux opérateurs n'est pas maximale mais lipschitzienne.

Voici les deux résultats dûs à Brézis [4]:

Soient A et B deux opérateurs maximaux monotones, l'opérateur $A+B$ est monotone. Mais, en général il n'est pas nécessairement maximal monotone, car son domaine $D(A+B)$ $(= D(A) \cap D(B))$ peut être vide.

Il existe un cas particulier avec une condition spéciale sur un des deux opérateurs où nous obtenons que l'opérateur-somme est maximal monotone :

Théorème 2.2.6 *(Brézis : [4]) Soient A un opérateur maximal monotone et B un opérateur monotone lipschitzien de X dans Y. Alors l'opérateur $A+B$ est maximal monotone.*

Puisque nous désirons poser comme hypothèses de départ que les deux opérateurs initiaux soient des opérateurs maximaux monotones, ces deux résultats précédents détectés dans [4] ne nous satisfont pas. Nous cherchons donc un autre théorème avec les hypothèses voulues et des conditions distinctes des précédentes et qui nous fournissent le résultat attendu, à savoir, la monotonie maximale pour l'opérateur-somme.
Nous l'obtenons avec une nouvelle condition posée par Rockafellar.

Le changement ne se fait pas uniquement au niveau de la condition, mais également pour l'espace de départ. Depuis le début de cette section, nous avons considéré que l'espace source des opérateurs était linéaire normé. Lorsque X est un espace linéaire normé réflexif, nous avons une condition de qualification de contrainte dûe à Rockafellar. L'application de cette condition nous donne le résultat suivant dont la preuve se trouve en [18] :

Théorème 2.2.7 *(Rockafellar : [18]) Soient A et B deux opérateurs maximaux monotones. Si $(Int(D(A))) \cap D(B) \neq \emptyset$, alors $A+B$ est un opérateur maximal monotone, et $\overline{D(A) \cap D(B)} = \overline{D(A)} \cap \overline{D(B)}$.*

Mais, cette condition sur l'intérieur de l'intersection des deux domaines des opérateurs A et B est une condition très forte. Dernièrement, Attouch, Riahi et Théra [1] ont proposé un résultat faisant intervenir une condition de qualification de contrainte plus faible, qu'ils utilisent dans le cas d'espaces de Banach réflexifs en dimensions finies. Le nouveau principe de dualité abstraite leur permet de prouver le résultat suivant comprenant la condition de qualification de contrainte plus faible.

Théorème 2.2.8 *(Attouch, Riahi et Théra) Soient $A, B : I\!R^n \rightrightarrows I\!R^n$ deux opérateurs maximaux monotones tels que*

$$\overline{D(A)} \cap \overline{D(B)} \neq \emptyset \tag{2.20}$$

et

$$\bigcup_{\lambda \geq 0} \lambda \left(D(A) - D(B)\right) \tag{2.21}$$

est un sous-espace linéaire fermé.
Alors, l'opérateur-somme $A+B$ est un opérateur maximal monotone.

Avant de commencer la démonstration de ce théorème, nous devons introduire quelques nouvelles définitions utiles et également quelques théorèmes intermédiaires que nous admettons.

Définition 2.2.9 *Si A est un opérateur maximal monotone, nous notons par f_∞^A la fonction de récession de A, définie par*

$$f_\infty^A(y) = \sup\left\{\langle y, x\rangle | x \in \overline{im(A)}\right\}$$

i.e., f_∞^A *est la fonction de suppport ou d'appui de l'ensemble convexe fermé* $\overline{im(A)}$.

Nous disons que l'opérateur maximal monotone A vérifie la condition de Brézis-Haraux si

$$(B-H) \qquad \forall x \in im(A), \forall y \in D(A), \sup_{(w,z)\in Gr(A)} \langle z-x, y-w\rangle < +\infty$$

Remarque 2.2.10 *La résolvante* $J_{\lambda A} = (I+\lambda A)^{-1}$ *satisfait toujours la condition de Brézis-Haraux, i.e.*

$$\forall x \in im(J_{\lambda A}),\ \forall y \in D(J_{\lambda A}), \sup_{(w,z)\in Gr(J_{\lambda A})} \langle z-x, y-w\rangle < +\infty$$

Notons par Φ *l'expression suivante*

$$\Phi = \sup_{(w,z)\in Gr(J_{\lambda A})} \langle z-x, y-w\rangle$$

Puisque $\forall x \in im(J_{\lambda A})$, *nous avons qu'il existe* $\tilde{x} \in D(J_{\lambda A})$ *tel que* $(\tilde{x}, x) \in Gr(J_{\lambda A})$, *et comme* $(w,z) \in Gr(J_{\lambda A})$, *nous obtenons*

$$\Phi = \sup_{w\in D(J_{\lambda A})} \langle J_{\lambda A}(w) - J_{\lambda A}(\tilde{x}), y-w\rangle$$

car $J_{\lambda A}$ *est univoque.*
Comme il est défini $\forall x \in im(J_{\lambda A})$, *en particulier nous pouvons prendre* $\tilde{x} = y$, *ainsi*

$$\begin{aligned}\Phi &= \sup_{w\in D(J_{\lambda A})} \langle J_{\lambda A}(w) - J_{\lambda A}(y), y-w\rangle \\ &= \sup_{w\in D(J_{\lambda A})} -\Big(\langle J_{\lambda A}(w) - J_{\lambda A}(y), w-y\rangle\Big) \\ &= -\inf_{w\in D(J_{\lambda A})} \langle J_{\lambda A}(w) - J_{\lambda A}(y), w-y\rangle\end{aligned}$$

Puisque la résolvante $J_{\lambda A}$ *est monotone, nous avons*

$$\langle J_{\lambda A}(w) - J_{\lambda A}(y), w-y\rangle \geq 0$$

et donc $\Phi \leq 0$, *c'est-à-dire qu'il est fini.*
La résolvante $J_{\lambda A}$ *satisfait donc bien toujours la condition de Brézis-Haraux.*

Nous avons également besoin de ces deux résultats suivants qui jouent un rôle important dans la démonstration du Théorème 2.2.8 :

Théorème 2.2.11 *(Brézis et Haraux : voir [5], Théorème 4) Soit* H *un espace de Hilbert et soient* $A, B : H \rightrightarrows H$ *deux opérateurs maximaux monotones tels que :*
(a) B *satisfait la condition de Brézis-Haraux;*
(b) $D(A) \subset D(B)$;

(c) $A+B$ est un opérateur maximal monotone.
Alors, $f_\infty^{A+B} = f_\infty^A + f_\infty^B$, i.e. la fonction de support de la somme de deux opérateurs maximaux monotones est équivalent à la somme des deux fonctions de support des deux opérateurs maximaux monotones.

Théorème 2.2.12 *(Attouch, Chbani et Moudafi: voir [1]) Soient H un espace de Hilbert, $A : H \rightrightarrows H$ un opérateur maximal monotone et f_∞^A sa fonction de récession.*
Supposons que

(i) Hypothèse de compacité: $\forall t_n \to +\infty, \forall v_n \to v$ faiblement avec $A(t_n v_n)$ borné, nous avons $v_n \to v$ fortement.

(ii) Condition de compatibilité:

(iia) $f_\infty^A \geq 0$

(iib) $Ker\, f_\infty^A$ est un sous-espace.

Alors, l'équation

$$0 \in Au$$

a au moins une solution.

Preuve du Théorème 2.2.8
Nous devons prouver que $A+B$ est un opérateur maximal monotone.
D'après le Théorème de Minty (voir le Théorème 1.6.11), cela revient à montrer que $I+A+B$ est surjectif.
Ce qui est équivalent à

$$\forall \beta \in H,\ \exists\ \alpha \in H \textit{ tels que } \beta \in \alpha + A\alpha + B\alpha \tag{2.22}$$

L'équation (2.22) est vérifiée pour tout β, donc en particulier pour $\beta\alpha$, l'équation (2.22) est donc équivalente à

$$\forall (\beta\alpha) \in H,\ \exists\ \alpha \in H \textit{ tels que } \beta\alpha \in \alpha + A\alpha + B\alpha$$

ou écrite plus simplement

$$\forall \beta \in H,\ \exists\ \alpha \in H \textit{ tels que } 0 \in (I+A)\alpha + (B-\beta)\alpha \tag{2.23}$$

Nous appliquons la définition (2.2.1) à l'équation (2.23), nous obtenons

$$\exists\ y \in H \textit{ tel que } 0 \in (I+A)^{-1}(y) - (B-\beta)^{-1}(-y) \tag{2.24}$$

en utilisant le changement de variable $y = (I+A)\alpha$.
Notons que, dans l'équation (2.24), l'opérateur $\tilde{A} = (I+A)^{-1}$ est un opérateur maximal monotone qui est une contraction (vérifiée par la définition de non-expansivité de l'opérateur résolvante) partout définie, et $\tilde{B}(.) = -(B-\beta)^{-1}(-.)$ est un opérateur maximal monotone

(voir Section 1.6).
Par le Théorème 2.2.6, le résultat $\tilde{C} = \tilde{A} + \tilde{B}$ est nécessairement un opérateur maximal monotone, puisque l'opérateur $\tilde{A}$ est une contraction donc lipschitienne. Nous vérifions ainsi dans notre cas la condition (iii) du Théorème 2.2.11.
Donc, le problème de la monotonie maximale de $A+B$ peut être remplacé de façon équivalente par une équation comprenant l'opérateur $\tilde{C}$ maximal monotone, ce qui donne le problème suivant

$$(PV) \qquad \textit{Trouver } y \in H \textit{ tel que } 0 \in \tilde{C}y$$

Remarquons que $im(\tilde{A}) = D(A)$ et $im(\tilde{B}) = -D(B)$.
Donc,

$$f_\infty^{\tilde{A}}(y) = \sup\left\{\langle y, u\rangle | u \in \overline{D(A)}\right\} \tag{2.25}$$

et

$$f_\infty^{\tilde{B}}(y) = \sup\left\{\langle y, u\rangle | u \in \overline{-D(B)}\right\} \tag{2.26}$$

Nous voyons que l'équation(2.24) nous fournit la condition (ii) du Théorème 2.2.11, à savoir $D(\tilde{B}) \subset D(\tilde{A})$.
Nous répondons donc bien aux conditions du Théorème 2.2.11 grâce à la remarque 2.2.10, qui nous assure que $\tilde{A}$ vérifie la condition de Brézis-Haraux, ce qui correspond à la condition (i) du Théorème 2.2.11.
Donc, par le Théorème 2.2.11, nous combinons les équations (2.25) et (2.26), et nous en déduisons

$$f_\infty^{\tilde{C}}(y) = f_\infty^{\tilde{A}+\tilde{B}}(y) = \sup\left\{\langle y, u - v\rangle | u \in \overline{D(A)}, v \in \overline{D(B)}\right\} \tag{2.27}$$

$f_\infty^{\tilde{C}}(y)$ est donc la fonction de support de $\overline{D(A)} - \overline{D(B)}$ i.e.

$$f_\infty^{\tilde{C}} = s(\overline{D(A)} - \overline{D(B)}; .) \tag{2.28}$$

Il nous reste à vérifier les conditions de compatibilité du Théorème 2.2.12, puisque l'hypothèse de compacité est toujours vérifiée en dimension finie.
Par la définition (2.28), la condition (iia) est équivalente à

$$s(\overline{D(A)} - \overline{D(B)}; .) \geq 0 \tag{2.29}$$

Nous pouvons remarquer que (2.29) est vraie si la condition

$$\overline{D(A)} \cap \overline{D(B)} \neq \emptyset \tag{2.30}$$

est satisfaite, i.e. la première hypothèse du Théorème 2.2.8 est vérifiée.
Il reste à vérifier que la deuxième hypothèse du Théorème 2.2.8 est satisfaite.
Puisque $\overline{D(A)} - \overline{D(B)} \subset \overline{D(A) - D(B)}$, nous avons que

$$f_\infty^{\tilde{C}}(y) = 0 \Longleftrightarrow \langle y, u\rangle \leq 0 \ \forall u \in \overline{D(A) - D(B)}$$

et donc $Kerf_{\infty}^{\tilde{C}}$ est un sous-espace linéaire si $\bigcup_{\lambda\geq 0}\lambda\,(D(A)-D(B))$ est un sous-espace fermé, et la preuve est terminée. □

Remarque 2.2.13 *L'hypothèse (2.21) est équivalente à*

$$0 \in sqi\,(D(A)-D(B)) \tag{2.31}$$

par la définition du quasi-intérieur au sens fort d'un sous-espace convexe $D(A)-D(B) \subset I\!\!R^n$. Donnons la définition générale du quasi-intérieur au sens fort d'un sous-espace convexe $C \subset X$, noté sqi C. Il est posé comme l'ensemble de tous les $x \in X$ pour lesquels $\bigcup_{\lambda\geq 0}\lambda(C-x)$ est un sous-espace fermé.

Grâce au Théorème 2.2.8, nous avons enfin la conclusion que la somme de deux opérateurs maximaux monotones est un opérateur maximal monotone, sous une hypothèse plus faible que celle posée par Rockafellar. Nous avons donc obtenu ce que nous voulions au sujet des opérateurs maximaux monotones en général. Si nous particularisons ces opérateurs, en considérant les opérateurs sous-différentiels, nous pouvons déduire du Théorème 2.2.8 une conclusion quant à la séparation du sous-différentiel d'une somme et un résultat concernant l'équivalence des "domaines".

Les opérateurs sous-différentiels

Nous savons par la Section 1.6 et par le Théorème de Rockafellar [18] que le sous-différentiel d'une fonction convexe, semi-continue inférieurement et propre est un opérateur maximal monotone sur H et par le Théorème de Brönsted-Rockafellar que pour toute fonction f du même type, $D(\partial f)$ est dense dans $dom f$.

Reprenons les hypothèses du Théorème 2.2.8 où nous posons $A = \partial f$ et $B = \partial g$, avec $f, g : H \longrightarrow I\!\!R \cup \{\infty\}$ deux fonctions convexes, semi-continues inférieurement et propres. Ces transformations appliquées au Théorème 2.2.8 nous donnent :

Théorème 2.2.14 *Soient $f, g : H \longrightarrow I\!\!R \cup \{\infty\}$ deux fonctions convexes, semi-continues inférieurement et propres telles que*

$$dom\ f \cap dom\ g \neq \emptyset \tag{2.32}$$

et

$$0 \in sqi\,(D(\partial f) - D(\partial g)) \tag{2.33}$$

Alors,

$$\partial f + \partial g = \partial(f+g)$$

Preuve
Nous savons par le Théorème 2.2.8 que $\partial f + \partial g$ et $\partial(f+g)$ sont des opérateurs maximaux monotones.
De plus, par le Théorème 1.5.7, nous avons que $\partial f + \partial g \subset \partial(f+g)$.

Puisque nous avons deux opérateurs maximaux monotones inclus l'un dans l'autre, ils ne peuvent qu'être égaux. □

Remarquons que le résultat précédent a recours à la différence des domaines des sous-différentiels ∂f et ∂g: $D(\partial f) - D(\partial g)$ plutôt que la différence habituellement utilisée des domaines effectifs de f et g: $dom\ f - dom\ g$.

Nous pouvons supposer que les ensembles $D(\partial f) - D(\partial g)$ et $dom\ f - dom\ g$ ont des propriétés semblables, nous présentons donc le résultat suivant :

Théorème 2.2.15 *Soient X un espace de Banach réflexif et $f, g : X \longrightarrow I\!R \cup \{\infty\}$ deux fonctions convexes, semi-continues inférieurement et propres sur X.*
Alors, le résultat suivant est vrai :

$$Int\,(dom\ f - dom\ g) = Int\,(D(\partial f) - D(\partial g)) \tag{2.34}$$

Preuve
Nous avons, dans [17], que

$$D(\partial f) \subset dom\ f$$

et donc également

$$D(\partial g) \subset dom\ g$$

Ce qui entraîne que

$$(D(\partial f) - D(\partial g)) \subset (dom\ f - dom\ g)$$

Nous en déduisons que

$$Int\,(D(\partial f) - D(\partial g)) \subset Int\,(dom\ f - dom\ g)$$

La première inclusion est vérifiée de façon évidente.
Il nous reste donc à montrer l'inclusion inverse, i.e.

$$Int\,(dom\ f - dom\ g) \subset Int\,(D(\partial f) - D(\partial g)) \tag{2.35}$$

Si nous prenons une fonction convexe, semi-continue inférieurement et propre sur X $h : X \longrightarrow I\!R \cup \{\infty\}$ à laquelle nous ajoutons une fonction continue convexe j, alors les domaines $dom\ h$ et $D(\partial h)$ restent inchangés, i.e.

$$dom\ (h + j) = dom\ h \quad et \quad D(\partial(h + j)) = D(\partial h)$$

En soustrayant de f et/ou de g une fonction minorante affine continue, nous pouvons supposer que f et g sont positifs.
En appliquant à f et/ou g la norme $\|.\|^2$, nous pouvons également supposer que f et/ou g est coercive.

Soit

$$x_0 \in Int\,(dom\ f - dom\ g)$$

En posant $f_{x_0}(u) = f(x_0 + u)$, nous avons que

$$0 \in Int\,(dom f_{x_0} - dom g)$$

Puisque $D(\partial f_{x_0}) = D(\partial f) - \{x_0\}$, nous pouvons observer que si

$$0 \in Int\,(D(\partial f_{x_0}) - D(\partial g))$$

alors

$$x_0 \in Int\,(D(\partial f) - D(\partial g))$$

et la preuve est terminée.

Supposons s.p.d.g. que $x_0 = 0$, et donc il reste à prouver que si

$$0 \in Int\,(dom\ f - dom\ g) \tag{2.36}$$

alors

$$0 \in Int\,(D(\partial f) - D(\partial g)) \tag{2.37}$$

Introduisons donc une fonction h définie par

$$h(x) = \inf_{u \in X} (f(x+u) + g(u))$$

Nous remarquons que

$$dom\ h = dom\ f - dom\ g \tag{2.38}$$

où, nous rappelons que, en général, $dom\ f = \{x \in X | f(x) < +\infty\}$

Par les hypothèses de coercivité sur f et/ou g, nous avons que h est une fonction convexe, propre, semi-continue inférieurement et positive. Le caractère semi-continu inférieurement suit de la coercivité de f et/ou g.
Nous avons aussi que $h(0) < \infty$, i.e.

$$\inf_{u \in X} (f(u) + g(u)) < \infty$$

ce qui est logique car

$$dom\ h = \{x \in X | \inf_{u \in X} (f(x+u) + g(u)) < +\infty\}$$

De plus, comme X est réflexif et que f et/ou g est coercive, pour tout $x \in X$, il existe un élément $u(x) \in X$ tel que

$$h(x) = f(x + u(x)) + g(u(x)) \tag{2.39}$$

Donc l'équation (2.36) est équivalente à

$$0 \in Int\ dom\ h \tag{2.40}$$

Pour toute fonction convexe, propre, semi-continue inférieurement h sur X, nous avons que $Int\ dom\ h = Int\ D(\partial h)$ (voir [17]).
Nous en déduisons donc que

$$0 \in Int\ D(\partial h) \tag{2.41}$$

Calculons maintenant le sous-gradient ∂h.
Supposons que $x \in D(\partial h)$, par définition, cela signifie qu'il existe $x^* \in X$ tel que pour tout $z \in X$

$$h(z) \geq h(x) + \langle x^*, z - x \rangle$$

i.e., pour tout $y, z \in X$

$$f(z+y) + g(y) \geq f(x + u(x)) + g(u(x)) + \langle x^*, z - x \rangle \tag{2.42}$$

En prenant $y = u(x)$ dans (2.42), nous obtenons

$$f(z + u(x)) \geq f(x + u(x)) + \langle x^*, z - x \rangle$$

dont nous déduisons que

$$x^* \in \partial f(x + u(x))$$

(qui est différent de l'ensemble vide), i.e.,

$$x + u(x) \in D(\partial f) \tag{2.43}$$

De même, en prenant $z + y = x + u(x)$ dans (2.42), nous obtenons

$$g(y) \geq g(u(x)) + \langle x^*, u(x) - y \rangle$$

ou plutôt

$$g(y) \geq g(u(x)) + \langle -x^*, y - u(x) \rangle$$

dont nous déduisons que

$$-x^* \in \partial g(u(x))$$

(qui est différent de l'ensemble vide), i.e.,

$$u(x) \in D(\partial g) \tag{2.44}$$

En combinant les équations (2.43) et (2.44), nous avons qu'il existe $u(x) \in X$ tel que

$$x = (x + u(x)) - u(x) \in D(\partial f) - D(\partial g)$$

i.e.,

$$D(\partial h) \subset D(\partial f) - D(\partial g) \tag{2.45}$$

En assemblant les équations (2.40), (2.45) et l'égalité suivante $Int\ dom\ h = Int\ D(\partial h)$, nous obtenons finalement l'inclusion $0 \in Int\,(D(\partial f) - D(\partial g))$ c'est-à-dire (2.37) et la preuve est donc terminée. □

Conclusions

Nous venons de présenter un nouveau principe de dualité dit : Dualité Abstraite, et ensuite nous l'avons appliqué à un problème soulevé précédemment, la dualité de Fenchel-Rockafellar. Ce principe est attrayant, il généralise un grand nombre de principes de dualité bien connus. Nous avons développé certains bénéfices et aussi des désavantages à l'application de ce nouveau principe.

Ce principe est très avantageux. En général, si le problème de départ (PV1) n'est pas bien posé, alors le problème dual (PV2) est mieux posé, et dans le cas où l'inclusion de départ est univoque, l'inclusion duale peut être multivoque.

Enfin, nous avons exposé quelques résultats qui sont améliorés grâce à l'apparition du nouveau principe général de dualité abstraite, au sujet de la somme d'opérateurs maximaux monotones généraux et sous-différentiels.

2.2.4 Exemple introductif au chapitre suivant

Soit le problème d'équilibre économique écrit sous la forme

$$0 \in \begin{pmatrix} -d(p) \\ f \end{pmatrix} + \begin{pmatrix} 0 & M \\ -M^* & 0 \end{pmatrix} \begin{pmatrix} p \\ y \end{pmatrix} + \partial I_P(p) \times \partial I_Y(y) \tag{2.46}$$

où, d est une fonction fortement non-linéaire de $I\!R^k$ vers $I\!R^k$, f est un vecteur constant de $I\!R^m$, M est une matrice $k \times m$ dont la transposée est M^*, et ∂I_P et ∂I_Y sont les opérateurs du cône normal associés avec deux ensembles convexes polyhédriques P et Y dans $I\!R^k$ et $I\!R^m$ respectivement.

Cela nous donne un problème à évaluer de dimension $k + m$, notre but est de trouver une transformation duale telle que le problème dual de (2.46) à résoudre soit de dimension moindre.

Introduisons la transformation linéaire S de $I\!R^{k+m}$ vers $I\!R^k$, définie par $S(x, y) = x$, alors nous pouvons réécrire (2.46) sous cette forme

$$0 \in S^*(-d)S(p, y) + A(p, y) \tag{2.47}$$

où, A est un opérateur maximal monotone de $I\!R^{k+m}$ vers $I\!R^{k+m}$ et donné par :

$$A(p, y) = \begin{pmatrix} 0 \\ f \end{pmatrix} + \begin{pmatrix} 0 & M \\ -M^* & 0 \end{pmatrix} \begin{pmatrix} p \\ y \end{pmatrix} + \partial I_P(p) \times \partial I_Y(y)$$

et S^* représente la transformation linéaire adjointe de S.

L'équation (2.47) est justifiée, S est défini de $I\!R^{k+m}$ dans $I\!R^k$, donc S^* est défini de $I\!R^k$ dans $I\!R^{k+m}$, et puisque

$$\langle S(x,y), z^* \rangle = \langle (x,y), S^*(z^*) \rangle$$

avec z^* défini sur $I\!R^k$.
Nous avons $\langle S(x,y), z^* \rangle = x^T z^*$, et donc

$$\langle (x,y), S^*(z^*) \rangle = x^T z^* = \begin{pmatrix} x^T & y^T \end{pmatrix} \begin{pmatrix} z^* \\ 0 \end{pmatrix}$$

Ainsi,

$$S^*(-d)S(p,y) = S^*(-d)p = S^*(-d(p)) = \begin{pmatrix} -d(p) \\ 0 \end{pmatrix}$$

La transformation linéaire adjointe S^* de la transformation linéaire S est définie comme suit

$$S^*(d) = \begin{pmatrix} d \\ 0 \end{pmatrix}$$

L'exemple présenté est purement économique. Il n'est pas semblable aux inclusions que nous avons utilisées dans les exposés précédents, il ne comprend plus deux opérateurs mais un opérateur et une composition d'une fonction et d'une transformation linéaire et de sa transformation adjointe, composition que nous énonçons au chapitre suivant.

Si nous pouvions appliquer au problème (2.47) un principe similaire à celui développé par Attouch et Théra, nous pourrions diminuer la dimension du problème et donc simplifier la résolution de ce problème.

Chapitre 3

Composition

Plusieurs auteurs, dont Stephen M. Robinson [15], se sont interrogés sur la composition des opérateurs et au sujet de leur monotonie maximalité. Un article récent de Robinson nous a introduit à ce problème. Nous avons découvert qu'en étendant le principe de dualité abstraite de Attouch et Théra [1], nous pouvons en déduire un principe qui permet de résoudre plus facilement une inclusion comprenant une composition d'opérateurs que nous appliquons à l'exemple économique présenté à la fin du chapitre 2. La deuxième partie de ce chapitre présentera l'évolution des connaissances au sujet de la composition des opérateurs maximaux monotones et de la monotonie et la maximalité de l'opérateur-composition.

3.1 Composition et dualité

Dans cette section, nous rappelons la définition de la composition comprenant entre autre des opérateurs. Ensuite, nous prenons un exemple intéressant, à l'aide duquel nous montrons l'intérêt de la recherche d'un nouveau principe. Puis, nous expliquons comment le principe général de dualité abstraite de Attouch et Théra pour des inclusions d'opérateurs peut être étendu aux problèmes variationnels comportant une composition d'opérateurs selon Robinson. Enfin nous l'appliquons au problème purement économique présenté au chapitre précedent et nous voyons en quoi le nouveau principe démontré facilite la résolution de ce problème.

3.1.1 Composition d'opérateurs

Comme nous l'avons dit plus haut, nous considérons maintenant les compositions d'opérateurs, il devient donc nécessaire de rappeler quelques notions importantes au sujet des compositions.

Revoyons d'abord la composition d'un opérateur et d'une application linéaire, rappelons que X et Y sont des espaces linéaires :

Définition 3.1.1 *Soit $A : X \rightrightarrows Y$ un opérateur, et soient $K : Y \longrightarrow Z$ et $L : Z \longrightarrow X$ deux applications linéaires, où Z est un espace de Hilbert.*

La composition $K \circ A$ est définie par :

$$(K \circ A)(x) \overset{\text{noté}}{=} K(A(x)) \overset{\text{déf}}{=} \bigcup_{y \in A(x)} Ky$$

De même, la composition $A \circ L$ est définie par :

$$(A \circ L)(z) \overset{\text{noté}}{=} A(Lz)$$

Considérons maintenant la composition de deux opérateurs multivoques, où X, Y et Z sont des espaces linéaires :

Définition 3.1.2 *Soient $A : X \rightrightarrows Y$ et $B : Y \rightrightarrows Z$ deux opérateurs multivoques.*
La composition $B \circ A$ est définie par :

$$(B \circ A)(x) \overset{\text{noté}}{=} (BA)(x) \overset{\text{déf}}{=} \{z \in Z \mid \exists\ y \in Ax \text{ avec } z \in By\}$$

Nous avons également annoncé que nous allions étendre le principe général de dualité abstraite présenté par Attouch et Théra qui implique une inclusion d'opérateurs. Il est intéressant de le rappeler ici, pour une compréhension plus rapide dans la suite de notre extension de ce principe.

Soient X et Y, deux espaces linéaires et soient A et C des opérateurs multivoques de X vers Y.
L'inclusion suivante

$$0 \in A(x) + C(x)$$

est vérifiée si et seulement si, l'inclusion duale

$$0 \in A^{-1}(y) - C^{-1}(-y)$$

est vérifiée.

Pourquoi Robinson s'intéresse à la composition?

Nous avons vu au premier chapitre les conditions sous lesquelles une somme d'opérateurs maximaux monotones conserve la monotonie et la maximalité. Dans ce chapitre, nous nous intéressons à des inclusions d'opérateurs comportant une somme et une composition avec des opérateurs maximaux monotones. Le problème de la somme étant résolu, il nous reste à montrer que la composition est maximale montone et nous pourrons affirmer que l'inclusion entière comporte un opérateur maximal monotone. Robinson [15] a donc développé une théorie et des conditions pour avoir la monotonie et la maximalité sur la composition d'un opérateur maximal monotone et d'une transformation linéaire et de sa transformation adjointe. Avec cette théorie, nous sommes sûrs d'avoir affaire à une somme d'opérateurs maximaux monotones, et donc nous pouvons appliquer les conditions vues au premier chapitre.

Avant de démontrer cette résolution du problème de la composition, nous présentons un résultat sur le problème dual équivalent à une inclusion comportant une composition, son application et son intérêt sur quelques exemples.

3.1.2 Exemple abstrait

Nous présentons maintenant un exemple dont la structure de base est très simple. Nous n'y faisons intervenir que des opérateurs dont nous ne donnons pas l'expression, nous travaillons donc sur un cas très général.

Soit l'inclusion variationnelle suivante :

$$0 \in A(x) + B(x)$$

où A et B sont des opérateurs multivoques de $I\!R^k$ vers $I\!R^n$ et donc $x \in I\!R^k$.

Supposons que la valeur de k est beaucoup plus élévée que la valeur de n, donc pour trouver le zéro de cette inclusion nous devons résoudre un problème (dit primal) de dimension k avec $k >>> n$.

Lorsque nous considérons son problème dual par le principe de dualité abstraite dû à Attouch et Théra, nous obtenons l'inclusion suivante :

$$0 \in A^{-1}(y) - B^{-1}(-y)$$

où $y \in I\!R^n$.
Le problème dual à résoudre est donc de difficulté moindre par rapport à son problème primal, nous avons un problème de dimension n. Il est donc intéressant de travailler avec le problème dual plutôt qu'avec le problème primal dans certains cas.

Considérons un problème primal où nous avons une composition d'opérateurs, et voyons ce qu'il en est lorsque nous rapprochons ce problème du principe de dualité abstraite défendu par Attouch et Théra.

Soit l'inclusion variationnelle suivante :

$$0 \in A(x) + BCD(x)$$

où A, B, C et D sont des opérateurs multivoques, A de $I\!R^k$ vers $I\!R^n$, B de $I\!R^m$ vers $I\!R^n$, C de $I\!R^l$ vers $I\!R^m$ et D de $I\!R^k$ vers $I\!R^l$ et $x \in I\!R^k$.
Si nous considérons également dans ce cas que la valeur de k est beaucoup plus importante

que les autres valeurs, i.e. $k >>> l, m, n$, le problème que nous devons résoudre est de dimension très élevée k.

La solution que nous espérons est qu'il existe un principe, proche de celui de dualité abstraite, qui permet d'obtenir un problème dual de dimension moindre, c'est ce que nous développons dans la partie qui suit.

3.1.3 Extension

Nous étendons maintenant le principe général de dualité abstraite de Attouch et Théra, faisant intervenir une somme d'opérateurs multivoques, à une composition d'opérateurs multivoques.

Pour rappel, le problème auquel nous appliquons le principe de dualité abstraite s'écrit comme suit :

Soient X et Y, deux espaces linéaires et soient A et B deux opérateurs multivoques de X vers Y. Nous avons l'inclusion suivante :

$$0 \in A(x) + B(x)$$

Pour cette extension, nous avons besoin d'introduire deux espaces linéaires en plus de X et Y, U et V, et donc nous avons aussi trois nouvelles multifonctions, $P : X \longrightarrow U$, $B : U \longrightarrow V$,et $Q : V \longrightarrow Y$.
Nous obtenons donc le problème primal composé suivant

$$0 \in A(x) + QBP(x) \tag{3.1}$$

Si nous appliquons le principe de dualité abstraite au problème (3.1) considéré dans cet ordre quant à la somme, nous obtenons

$$0 \in A^{-1}(z) - (QBP)^{-1}(-z)$$

où $z \in Y$, que nous résumons par le principe de dualité suivant :

Théorème 3.1.3 *Le problème variationnel*

$$0 \in A(x) + QBP(x)$$

est équivalent au problème variationnel

$$0 \in (-P)A^{-1}(-Q)(z) + B^{-1}(z) \tag{3.2}$$

Preuve

$\boxed{\Longrightarrow}$ La résolution de (3.1) implique qu'il existe u tel que

$$0 = u - u \in A(x) - \Big(- QBP(x)\Big)$$

d'où il existe

$$u \in A(x) \cap \Big(- QBP(x)\Big)$$

Si $u \in A(x)$, alors $x \in A^{-1}(u)$. (1ère condition)

Si $u \in \Big(-QBP(x)\Big)$, alors il existe $v \in P(x)$ tel que $u \in \Big(-QB(v)\Big)$ et il existe $z \in B(v)$ tel que $u \in \Big(- Q(z)\Big)$.

Donc, si $u \in \Big(- QBP(x)\Big)$, alors il existe v et z tels que (2ème condition)

$$u \in \Big(- Q(z)\Big), \ v \in P(x), \ z \in B(v),$$

Ainsi, par la 2ème condition,

$$v \in B^{-1}(z)$$

et, par les deux conditions,

$$-v \in \Big(- P(x)\Big) \implies -v \in (-P)A^{-1}(u) \implies -v \in (-P)A^{-1}(-Q)(z)$$

Donc,

$$v - v = 0 \in B^{-1}(z) + (-P)A^{-1}(-Q)(z)$$

i.e. z résout (3.2).

$\boxed{\Longleftarrow}$ Par symétrie, la résolution de (3.2) implique qu'il existe u tel que

$$u - u \in (-P)A^{-1}(-Q)(z) - \Big(- B^{-1}(z)\Big)$$

d'où il existe

$$u \in (-P)A^{-1}(-Q)(z) \cap (-B^{-1}(z))$$

Si $u \in \Big(- B^{-1}(z)\Big)$, alors $-u \in B^{-1}(z)$, et donc $z \in B(-u)$. (1ère condition)

Si $u \in (-P)A^{-1}(-Q)(z)$, alors il existe $v \in \Big(- Q(z)\Big)$ tel que $u \in (-P)A^{-1}(v)$ et il existe $x \in A^{-1}(v)$ tel que $u \in (-P)(x)$.

Donc, si $u \in (-P)A^{-1}(-Q)(z)$, alors il existe v et x tels que (2ème condition)

$$u \in (-P)(x), \ v \in \Big(- Q(z)\Big), \ x \in A^{-1}(v)$$

Ainsi, par la 2ème condition,

$$v \in A(x)$$

et, par les deux conditions,

$$-v \in Q(z) \implies -v \in QB(-u) \implies -v \in QBP(x)$$

Donc,

$$v - v = 0 \in A(x) + QBP(x)$$

et ainsi x résout (3.1). □

Remarque 3.1.4 *Cas particulier*
Si nous considérons le Théorème 3.1.3 avec les changements suivants : $U = X$, $V = Y$ et $B = C$, et si nous posons que les multifonctions P et Q sont des applications identités de X et Y respectivement, nous retrouvons le problème variationnel de la dualité développé par Attouch et Théra, voir la Définition 2.2.1. Ce qui permet de dire que nous avons bien étendu le principe de dualité abstraite de la somme d'opérateurs sur la composition d'opérateurs, puisque en posant certaines conditions qui font disparaître la composition (nous faisons intervenir des applications identités), nous revenons à ce principe.

Remarque 3.1.5 *Le dual du dual est le primal.*
Nous pouvons, comme au chapitre premier pour le principe de dualité abstraite, vérifier que le dual du problème dual est le problème primal.

Soit le problème primal général :

$$0 \in A(x) + QBP(x)$$

Appliquons-y le Théorème 3.1.3, nous obtenons :

$$0 \in (-P)A^{-1}(-Q)(z) + B^{-1}(z)$$

Ce problème répond aux hypothèses pour appliquer une deuxième fois le théorème. Nous le réécrivons sous la même forme pour pouvoir l'appliquer plus rapidement. Donc, nous écrivons

$$0 \in B^{-1}(z) + (-P)A^{-1}(-Q)(z)$$

et en appliquant le théorème, nous obtenons

$$0 \in (A^{-1})^{-1}(x) + (-(-Q))(B^{-1})^{-1}(-(-P))(x)$$

i.e.,

$$0 \in A(x) + QBP(x)$$

ce qui nous rend le problème primal, ce que nous espérions.

3.1.4 Reprise des exemples

Reprenons l'exemple abstrait vu dans ce chapitre et que nous voulions dualiser pour obtenir un problème de dimension moindre :

Soit le problème variationnel suivant :

$$0 \in A(x) + BCD(x)$$

où A, B, C et D sont des opérateurs multivoques, A de $I\!R^k$ vers $I\!R^n$, B de $I\!R^m$ vers $I\!R^n$, C de $I\!R^l$ vers $I\!R^m$ et D de $I\!R^k$ vers $I\!R^l$ et $x \in I\!R^k$.
Rappelons également que nous considérons le cas où $k >>> l, m$ et n.

Appliquons le nouveau principe de dualité décrit au Théorème 3.1.3 à ce problème, nous obtenons l'inclusion équivalente suivante :

$$0 \in (-D)A^{-1}(-B)(z) + C^{-1}(z) \qquad z \in I\!R^m$$

Nous voyons donc qu'en appliquant le nouveau principe défini au Théorème 3.1.3, nous obtenons un problème plus facile à résoudre avec une dimension m qui est beaucoup plus petite que la dimension du problème primal ($m <<< k$).

Considérons maintenant un exemple un peu moins abstrait. Cet exemple a été présenté à la fin du chapitre premier, et nous cherchons un principe de dualité qui peut aboutir à un problème équivalent plus simple à résoudre.

Reprenons donc le problème purement économique que nous examinons à partir du moment où nous avons opéré quelques simplifications, i.e.

$$0 \in A(p, y) + S^*(-d)S(p, y) \tag{3.3}$$

où, S est la transformation linéaire de $I\!R^{k+m}$ vers $I\!R^k$, définie par $S(x, y) = x$, S^* représente la transformation linéaire adjointe de S définie par

$$S^*(d) = \begin{pmatrix} d \\ 0 \end{pmatrix}$$

et A est un opérateur maximal monotone de $I\!R^{k+m}$ vers $I\!R^{k+m}$ donné par :

$$A(p, y) = \begin{pmatrix} 0 \\ f \end{pmatrix} + \begin{pmatrix} 0 & M \\ -M^* & 0 \end{pmatrix} \begin{pmatrix} p \\ y \end{pmatrix} + \partial I_P(p) \times \partial I_Y(y)$$

avec, d une fonction fortement non-linéaire de $I\!R^k$ vers $I\!R^k$, f un vecteur constant de $I\!R^m$, M une matrice $k \times m$ dont la transposée est M^*, et ∂I_P et ∂I_Y les opérateurs du cône normal associé aux deux ensembles convexes polyhédriques P et Y dans $I\!R^k$ et $I\!R^m$ respectivement.

Appliquons-y la transformation de dualité que nous venons de démontrer, nous obtenons

$$0 \in (-d)^{-1}(q) + (-S)A^{-1}(-S^*)(q) \tag{3.4}$$

Pour simplifier le résultat et faciliter sa lecture, nous appliquons le principe général de dualité abstraite à l'équation (3.4),

$$0 \in (-d)(p) - (-S^*)^{-1}A(-S)^{-1}(-p)$$

Nous éliminons les signes négatifs dans les limites du possible de la façon suivante :
Si nous notons

$$q \in (-S^*)^{-1}A(-S)^{-1}(-p)$$

alors

$$\begin{array}{rcl}
(-S^*)(q) & \in & A(-S)^{-1}(-p) \\
& \Updownarrow & \\
S^*(-q) & \in & A(-S)^{-1}(-p) \\
& \Updownarrow & \\
A^{-1}S^*(-q) & \in & (-S)^{-1}(-p) \\
& \Updownarrow & \\
(-p) & \in & (-S)A^{-1}S^*(-q) \\
& \Updownarrow & \\
p & \in & -\Big((-S)A^{-1}S^*(-q)\Big) \\
& \Updownarrow & \\
p & \in & SA^{-1}S^*(-q) \\
& \Updownarrow & \\
(-q) & \in & (S^*)^{-1}AS^{-1}(p)
\end{array}$$

obtenu par le fait que S et S^* sont des transformations linéaires, et donc

$$0 \in -d(p) + (S^*)^{-1}AS^{-1}(p)$$

ou plus simplement encore,

$$d(p) \in (S^*)^{-1}AS^{-1}(p) \tag{3.5}$$

où, S^{-1} est l'inverse de S dans le sens des multifonctions.

Remarque 3.1.6 *Cette même démarche peut être faite plus directement, nous pouvons simplifier comme ceci :*

$$(-S)^{-1}(-p) = S^{-1}(p) \quad \textit{et} \quad (-S^*)^{-1} = -(S^*)^{-1}$$

et nous obtenons le même résultat final, à savoir :

$$0 \in -d(p) + (S^*)^{-1}AS^{-1}(p)$$

La solution du problème dual (3.5) produit à l'aide de quelques calculs la solution du problème primal (3.3). Donc, en passant par le problème dual, nous réduisons la dimension du problème et nous le résolvons.

Nous sommes passé d'un problème primal à $k + m$ dimensions à son problème dual à k dimensions en appliquant le nouveau principe de dualité décrit au Théorème 3.1.3. Avant l'application du principe, nous devions chercher les solutions en (p, y) appartenant à $I\!R^{k+m}$, maintenant, nous cherchons les solutions en p appartenant à $I\!R^k$.

Si les dimensions du problème étaient $k = 14$ et $m = 26$, le principe de dualité (3.1.3) transforme un problème à 40 variables en un problème à 14 variables.

Dans [16], nous avons un exemple d'application de ce principe avec un modèle d'équilibre de Walras, où nous observons que le temps de calcul est réduit de 98% grâce à l'application de ce principe.

Ce dernier problème comportait un opérateur maximal monotone et une certaine transformation linéaire S et son application adjointe S^*. Ce cas est très intéressant, il est l'objet essentiel de la section suivante et il rassemble un grand nombre de modèles de dualité.

Conclusions
Nous avons obtenu un principe de dualité s'appliquant aux inclusions comportant une composition d'opérateurs, c'est le Théorème 3.1.3. Nous l'avons démontré et appliqué aux deux exemples présentés précédemment.

Ce principe peut s'appliquer deux fois à un problème et nous réobtenons le problème de départ. De plus, nous pouvons dire que c'est une généralisation du principe de dualité présenté au chapitre précédent, le principe de dualité abstraite.

Ce nouveau principe comporte les mêmes inconvénients techniques, le calcul de la fonction conjuguée de la bonne fonction de perturbation n'est plus à exécuter, mais nous devons évaluer la fonction inverse de certains opérateurs, et ce n'est pas toujours une chose aisée.

3.2 Monotonie et maximalité

Nous nous intéressons maintenant à ce cas particulier présenté à la section précédente. La question est de savoir si cette composition d'un opérateur maximal monotone et de cette transformation linéaire et de sa transformation linéaire adjointe est encore un opérateur maximal monotone. Dans cette section, nous présentons quelques cas particuliers de cette composition. Ensuite, nous nous posons la question de savoir sous quelles conditions une telle composition est maximale monotone. Dans un premier temps, nous présentons les différents résultats connus sur le problème de la monotonie de la composition. Dans un second temps, nous présentons une théorie qui prouve que, sous une simple condition, la propriété de monotonie et maximalité est maintenue. Cette théorie comporte plusieurs résultats intermédiaires démontrés avant d'obtenir l'issue finale.

3.2.1 Développement d'un cas particulier

Comme prévu, reprenons le cas présenté à la section précédente, ce cas nous donne des applications numériques et des propriétés spéciales et différentes du Théorème 3.1.3. Dans la suite, nous en présentons quelques exemples particuliers.

Notre problème variationnel principal se présente comme suit

$$0 \in A(x) + S^* BS(x)$$

où, en général, A et B sont des opérateurs maximaux monotones, et S et S^* sont des transformations linéaires, la seconde étant l'application adjointe de la première.
Son problème variationnel dual équivalent est

$$0 \in (-S)A^{-1}(-S^*)(z) + B^{-1}(z)$$

Nous reprendrons toujours ces expressions de base sur lesquelles nous effectuons quelques changements ou quelques ajouts pour obtenir à chaque fois un exemple différent. Nous présentons en général à chaque fois le problème primal et son problème dual.

Remarque 3.2.1 *A partir d'ici, nous identifions les espaces primals avec leurs espaces duals. Cette condition nous permet à l'avenir des simplifications d'écriture.*

Voici notre premier exemple :

Supposons que H et K sont des espaces de Hilbert réels.
Soit A un opérateur maximal monotone de H vers H et soit S une transformation linéaire continue de H vers K et sa transformation linéaire adjointe S^*. Si B est une multifonction de K vers K (éventuellement maximale monotone), alors nous écrivons les inclusions du Théorème 3.1.3 comme suit

$$0 \in A(x) + S^* BS(x) \tag{3.6}$$

et

$$0 \in (-S)A^{-1}(-S^*)(z) + B^{-1}(z) \tag{3.7}$$

Si nous considérons le cas où les espaces sont identiques : $H = K$, la transformation linéaire S est la transformation identité : $S = I$, l'opérateur A est l'opérateur du cône normal associé à un sous-ensemble convexe fermé C de K : $A = \partial I_C = N_C$, et l'opérateur B est univoque.

Les inclusions (3.6) et (3.7) deviennent donc

$$0 \in N_C(x) + B(x) \tag{3.8}$$

$$0 \in (N_C)^{-1}(z) + B^{-1}(z)$$

L'inclusion (3.8) entraîne que

$$B(x) \in (-N_C(x))$$

Donc, par la définition du cône normal, nous obtenons le problème d'inéquation variationnelle suivant

$$Trouver \ \ x \in C \ \ t.q. \ \ \langle B(x), c - x \rangle \geq 0 \quad \forall \, c \in C$$

Les problèmes de ce type se retrouvent souvent dans beaucoup d'exemples de mathématiques appliquées.

Pour l'exemple suivant, nous supposons cette fois que les espaces de Hilbert réels H et K sont différents, et que les opérateurs A et B sont des applications sous-différentielles associées à des fonctions convexes, propres, semi-continues inférieurement f et g, respectivement, l'inclusion (3.6) devient

$$0 \in \partial f(x) + S^* \partial g S(x)$$

et l'inclusion (3.7) devient

$$0 \in (-S)(\partial f)^{-1}(-S^*)(z) + (\partial g)^{-1}(z)$$

Ces structures sont rencontrées assez souvent également dans différents cadres d'application.

L'exemple plus important découlant du précédent est peut être le cas suivant, où nous introduisons en plus un autre espace de Hilbert réel J.

Nous prenons pour l'opérateur A l'opérateur zéro sur l'espace $H = J \times K$, i.e., $A = 0 : H \rightrightarrows H$, et l'opérateur B est l'opérateur sous-différentiel associé à la fonction convexe, propre, semi-continue inférieurement g sur H, i.e. $B = \partial g : H \rightrightarrows H$.
Si, de plus, nous remplaçons la transformation linéaire S par la transformation linéaire $L : J \longrightarrow H = J \times K$, donnée par $L(x) = (x, 0)$, alors l'inclusion (3.6) devient

$$0 \in L^*(\partial g)L(x) \tag{3.9}$$

où, L^* est la transformation linéaire adjointe de L de $H = J \times K$ vers J et est définie comme la projection sur le premier espace J, i.e. $L^*(y, z) = y$ avec $(y, z) \in H = J \times K$. Nous remarquons qu'il n'y a pas de deuxième terme, il n'est pas explicité ici puisque l'opérateur zéro est sous-entendu.
Et l'inclusion (3.7) devient

$$0 \in (-L)(0^{-1})(-L^*)(y, z) + (\partial g)^{-1}(y, z) \tag{3.10}$$

L'inclusion (3.10) peut s'écrire aussi

$$0 \in (-L)(0^{-1})(-L^*)(y, z) + \partial g^*(y, z)$$

par une propriété bien connue des inverses des opérateurs sous-différentiels (voir Corollaire 1.5.6).

Remarque 3.2.2 *L'opérateur nul,* $0 : H \rightrightarrows H$, *qui à tout* $x \in H$ *fait correspondre* $0 \in H$, *a pour opérateur inverse, l'opérateur* $0^{-1} : H \rightrightarrows H$ *tel que, si* $z = 0$, *alors l'image par* 0^{-1}

est l'espace tout entier H, et si $z \neq 0$, alors l'image est vide. Nous voyons donc que c'est l'opérateur du cône normal $N_{\{0\}}$ sur H, c'est-à-dire par la Section 1.5 le sous-différentiel $\partial I_{\{0\}}$ de la fonction indicatrice de $\{0\}$.

Mosco [13] considérait l'exemple où la transformation linéaire est la transformation identité : $S = I$, A est un opérateur maximal monotone, B est l'application sous-différentielle associée à la fonction convexe, semi-continue inférieurement g : $B = \partial g$ et donc $B^{-1} = (\partial g)^{-1} = \partial g^*$, i.e. B^{-1} est l'application sous-différentielle associée à la fonction convexe conjugée g^* de g. Donc, le principe de dualité est dans ce cas :
L'inclusion

$$0 \in A(x) + \partial g(x)$$

est équivalente à l'inclusion

$$0 \in A^{-1}(z) + \partial g^*(z)$$

obtenu par le Théorème 3.1.3.

Plus tard, Gabay [8] étendait la dualité présentée par Mosco en posant que H et K peuvent être différents, que S est un opérateur linéaire continu de H vers K, A est un opérateur maximal monotone de H vers H, et l'opérateur B est encore considéré comme l'application sous-différentielle de K vers K associée à une fonction convexe, semi-continue inférieurement g.
Le principe de dualité est que, l'inclusion

$$0 \in A(x) + S^*(\partial g)S(x)$$

est équivalente à

$$0 \in (-S)A^{-1}(-S^*)(z) + \partial g^*(z)$$

Plus récemment, Fukushima [7] a étudié l'exemple lorsque A et B sont des opérateurs maximaux monotones généraux. Pour résoudre ces inclusions, il a utilisé les méthodes de séparation, mais pour pouvoir l'appliquer correctement, il avait besoin de l'assurance que les deux opérateurs soient maximaux monotones.

Il devient évident qu'il faut chercher des cas particuliers dont nous pouvons généraliser les conditions, les hypothèses afin de prouver qu'une composition comportant un opérateur maximal monotone est maximale monotone. Dans la suite, nous présentons quelques cas particuliers et ultérieurs à la théorie développée par Robinson.

3.2.2 Evolution de la monotonie et la maximalité

Nous rappelons ici les différentes connaissances au sujet de la monotonie et de la maximalité de la compositon d'une transformation linéaire et d'un opérateur maximal monotone pour un des exemples vu précédemment, et également les conditions proposées par Gabay et Fukushima pour résoudre ce problème.

Nous savons que, sous une simple hypothèse, ici : $im\ S \cap Int\ D(\partial f) \neq \emptyset$, la composition d'un opérateur linéaire continu S, d'un opérateur sous-différentiel d'une fonction convexe, propre, semi-continue inférieurement f, et de l'opérateur adjoint de S, S^*, i.e. la composition $S^*(\partial f)S$, est l'application sous-différentiel de la fonction convexe, propre, semi-continue inférieurement $f \circ S$: $S^*(\partial f)S = \partial(f \circ S)$, et nous avons que cette composition est maximale monotone.

Donc, si nous reprenons l'exemple proposé par Gabay, nous obtenons bien une somme d'opérateurs maximaux monotones et pour déduire la monotonie et la maximalité de la somme, nous avons résolu le problème au chapitre précédent. Dans le cas où les opérateurs A et S^*BS ne sont pas univoques (mais pas nécessairement maximaux monotones), il existe des méthodes d'application variées et disponibles, par exemple, pour résoudre le problème purement économique présenté plus haut, la méthode utilisée est appelée méthode d'homotopie (voir [16]).

Notre problème de savoir si la composition est maximale monotone ne se pose pas que dans le cas de l'inclusion primale, mais également dans l'inclusion duale.
Quand l'expression $(-S)A^{-1}(-S^*)$ est-elle aussi un opérateur maximal monotone?

Nous avons donc montré que, sous certaines conditions, pour le cas où les opérateurs A et B sont des opérateurs sous-différentiels, la monotonie et la maximalité de la composition est vérifiée facilement. Mais le problème qui nous intéresse ici est beaucoup plus général.

Nous considérons le problème où nous possèdons un opérateur maximal monotone T de H vers H, une transformation linéaire continue L de K vers H avec H et K deux espaces de Hilbert réels. Il est évident que la composition L^*TL est monotone, mais que pouvons-nous dire au sujet de sa maximalité?

Voici un exemple qui montre que même lorsque l'opérateur T est le sous-différentiel d'une fonction convexe, la composition L^*TL n'est pas nécessairement maximal monotone. Prenons l'opérateur T comme le sous-différentiel de la fonction indicatrice associée à l'ensemble des points $(x_1, x_2) \in I\!R^2$ tels que $x_2 \geq x_1^2$ et L est la transformation linéaire de $I\!R$ vers $I\!R^2$ définie par $L(x) = (x, 0)$. Donc, pour y appliquer l'opérateur T, nous devons répondre à la condition suivante : $0 \geq x^2$, ce qui entraîne que $x = 0$, et le graphe de L^*TL consiste en le seul point $(0,0)$, qui n'est évidemment pas maximal monotone.

Certains auteurs ont proposés des conditions, comme nous allons le voir brièvement, elles sont assez restrictives, d'où l'intérêt de développer une autre théorie plus générale.

Voici les conditions que Gabay [8] a prouvé :

Théorème 3.2.3 *L^*TL est maximal monotone si*
(a) T est maximal monotone, et
(b) soit T^{-1} est fortement monotone, soit LL^* est un isomorphisme de H.

La première condition implique que T est Lipschitzienne et donc, en particulier, à valeur unique. La seconde pose le problème qu'elle ne peut pas être vérifiée simultanément pour les inclusions primale et duale, sauf dans le cas où H et K sont isomorphes.

Fukushima [7] a utilisé la seconde alternative de la condition (b) dans son analyse.

Nous avons donc besoin de conditions qui restent valables aussi bien pour le primal que pour le dual, c'est ce qui fait l'objet de la partie suivante.

3.2.3 Théorie sur la maximalité

Dans cette section, nous supposons que H et K sont deux espaces de Hilbert réels, que T est un opérateur maximal monotone de H vers H, et que L est une transformation linéaire continue de K vers H. Notre principale question est de déterminer les conditions sous lesquelles la composition L^*TL est un opérateur maximal monotone, nous savons déjà que cette composition est monotone.

Si nous reprenons le cas particulier où $T = \partial f$, avec f une fonction convexe, propre, semi-continue inférieurement, nous trouvons facilement une condition. Comme nous savons que, si l'image de L imL intersecte l'intérieur du domaine du sous-différentiel de f $Int\ D(\partial f)$, alors la composée $L^*(\partial f)L$ est le sous-différentiel de la fonction convexe propre semi-continue inférieurement $f \circ L$, et est par conséquent maximal monotone (voir Section 1.6), i.e.,

$$im\ L\ \cap\ Int\ D(\partial f) \neq \emptyset \Longrightarrow L^*(\partial f)L = \partial(f \circ L)$$

Cette condition pour le cas des sous-différentiels, nous pouvons la généraliser et penser que la condition nécessaire pour que la composée L^*TL soit un opérateur maximal monotone est que l'image de L imL intersecte l'intérieur du domaine de l'opérateur maximal monotone T $Int\ D(T)$. Ceci n'est qu'une partie de ce que nous prouvons dans la suite, avec des conditions supplémentaires.

Nous avons remarqué dans notre étude que les résultats que nous présentons et démontrons pour certains sont emboîtés, c'est-à-dire que les théorèmes sont à tour de rôle des cas particuliers des précédents.

Donc, la maximalité de L^*TL dépend des sous-espaces $im\ L$ et $im\ L^*$, et d'aucun autre aspect de la transformation linéaire comme nous aurions pu le penser.

Voici un premier résultat que nous démontrons :

Théorème 3.2.4 *Supposons que H et K sont des espaces de Hilbert réels et que L est une transformation linéaire continue de K vers H telle que $im\ L^*$ est fermée dans K. Soit T un opérateur maximal monotone de H vers H. Si $T + \partial I_{imL}$ est maximal monotone, où I_{imL} est* *la fonction indicatrice de im L, alors* $M = L^*TL$ *est un opérateur maximal monotone de* K *vers* K.

Preuve
M est une multifonction de K vers K et un calcul simple montre qu'elle est monotone. Nous devons donc prouver que

$$\forall a, b \in K \quad \langle a - b, M(a) - M(b) \rangle \geq 0$$

Donc,

$$\langle a - b, L^*TL(a) - L^*TL(b) \rangle$$

est équivalent à

$$\langle L(a) - L(b), TL(a) - TL(b) \rangle$$

car L et L^* sont linéaires.
Et comme $L(a), L(b) \in H$, alors par la montonie de T, nous avons

$$\langle L(a) - L(b), T(L(a)) - T(L(b)) \rangle \geq 0$$

Supposons que (a, b) est un point de $K \times K$ tel que pour tout $(u, v) \in Gr(M)$, nous avons $\langle u - a, v - b \rangle \geq 0$. Nous obtiendrons la maximalité de la multifonction M si nous montrons que $(a, b) \in Gr(M)$.

Nous savons que $im\ L^*$ est fermée dans K, donc $im\ L^*$ est un sous-espace fermé de K. Nous écrivons b comme $b = L^*c + s$, où s est un élément du noyau $Ker\ L$ de L, i.e. $L(s) = 0$. Nous savons de plus, par hypothèse, que $im\ L$ rencontre $D(T)$.

Prenons $p_0 \in K$ tel que $Lp_0 \in D(T)$ et $t_0 \in T(Lp_0)$.
Pour un $\lambda \in I\!R$, nous définissons $p_\lambda = p_0 + \lambda s$, nous voyons que

$$L^*TL(p_\lambda) = L^*TL(p_0 + \lambda s) = L^*TL(p_0) + L^*TL(\lambda s) = L^*TL(p_0) + L^*T(0) = L^*TL(p_0) = L^*t_0$$

Donc, pour chaque réel λ, la paire (p_λ, L^*t_0) appartient au $Gr(M)$.

Donc, nous pouvons calculer

$$\begin{aligned}
0 &\leq \langle p_\lambda - a, L^*t_0 - b \rangle \\
&= \langle p_0 + \lambda s - a, L^*t_0 - b \rangle \\
&= \langle p_0 - a, L^*t_0 - b \rangle + \lambda \langle s, L^*t_0 - L^*c - s \rangle \\
&= \langle p_0 - a, L^*t_0 - b \rangle + \lambda \langle s, L^*(t_0 - c) - s \rangle \\
&= \langle p_0 - a, L^*t_0 - b \rangle + \lambda \langle s, L^*(t_0 - c) \rangle - \lambda \langle s, s \rangle \\
&= \langle p_0 - a, L^*t_0 - b \rangle + \lambda \langle Ls, t_0 - c \rangle - \lambda \|s\|^2 \\
&= \langle p_0 - a, L^*t_0 - b \rangle - \lambda \|s\|^2
\end{aligned}$$

Comme λ a été choisi arbitrairement, donc nous pouvons avoir $s = 0$, et donc $b = L^*c$.

Nous avons comme hypothèse que l'opérateur $T + \partial I_{imL}$ est maximal monotone.
Soit $(q, r) \in Gr(P = T + \partial I_{imL})$, alors $q = Lh$ pour certains h et $r = d + e$ avec $e \in Ker\ L^*$ et $(Lh, d) \in T$, ce qui est justifié car :

$$g \in P(q) = T(q) + \partial I_{imL}(q) \tag{3.11}$$

Soit $e \in \partial I_{imL}$, par la Section 1.5, nous savons que $\partial I_{imL}(q) = N_{imL}(q)$, le cône normal de q dans $im\ L$. Or, nous avons que $(imL)^\perp = Ker\ L^*$, et donc $e \in Ker\ L^*$.

Reprenons l'autre partie de l'expression (3.11), soit $d \in T(q)$ et donc $(q, d) \in T$.
D'où, $q \in D(T)$, nous savons par un résultat précedent que $Lp_0 \in D(T)$, cela signifie qu'il existe au moins un $h \in H$ tel que $q = Lh \in D(T)$. Donc, $(q, d) = (Lh, d) \in T$.

En résumé, $L^*d \in L^*TL(h) = L^*T(q)$ et par conséquent $(h, L^*d) \in Gr(M = L^*TL)$.

Par la définition de la monotonie des opérateurs, ceci implique que

$$\begin{aligned} 0 &\leq \langle h - a, L^*d - L^*c\rangle \\ &= \langle h - a, L^*(r - e) - L^*c\rangle \\ &= \langle h - a, L^*r - L^*e - L^*c\rangle \\ &= \langle h - a, L^*(r - c)\rangle \\ &= \langle L(h - a), r - c\rangle \\ &= \langle f - La, r - c\rangle \end{aligned}$$

Puisque $(q, r) \in Gr(P)$, nous avons par la maximalité de P que $(La, c) \in Gr(P)$.

Comme T est multivoque et comme $M = L^*TL$, alors il existe $c, c' \in TLa = Ta'$, avec $La = a'$ car L est univoque, tels que $L^*c = L^*c'$, et $(La, c), (La, c') \in Gr(T)$.
D'où, comme $(a, L^*c') \in M = L^*TL$, nous obtenons que $(a, L^*c') = (a, L^*c) = (a, b) \in Gr(M)$ car nous avons calculé que $L^*c = b$ dans le début de la preuve. □

Nous pouvons développer deux critères employant $im\ L$, impliquant chacun les hypothèses du Théorème 3.2.4, et donc la maximalité de L^*TL.
D'abord, nous voyons un résultat pour le cas où la dimension de H peut être infinie.

Proposition 3.2.5 *Soient H un espace de Hilbert réel et T un opérateur maximal monotone de H vers H. Soit N un sous-espace fermé de H. Si, soit N rencontre $Int\ D(T)$, soit $N = H$, alors $T + \partial I_N$ est maximal monotone.*

Preuve

- Si $N = H$, alors $\partial I_N = N^\perp$ est l'opérateur zéro sur tout H et il n'y a plus rien à prouver.

- Si N rencontre $Int\ D(T)$, alors puisque $D(\partial I_N) = N$, par le théorème de Rockafellar (voir Théorème 2.2.7), nous avons que $T + \partial I_N$ est maximal monotone. □

En combinant la Proposition 3.2.5 avec le Théorème 3.2.4, nous obtenons une condition suffisante pour la maximalité de L^*TL dans un espace de Hilbert.

Théorème 3.2.6 *Supposons que H et K sont des espaces de Hilbert réels et que L est une transformation linéaire continue de K vers H telle que $im\ L$ est fermée dans H et $im\ L^*$ est fermée dans K. Soit T un opérateur maximal monotone de H vers H. Si, soit $im\ L$ rencontre $Int\ D(T)$, soit $im\ L = H$, alors $M = L^*TL$ est un opérateur maximal monotone de K vers K.*

Preuve
Les deux hypothèses sous le 'si' entraîne par la Proposition 3.2.5 que $T + \partial I_{imL}$ est un opérateur maximal monotone, ce qui amène par le Théorème 3.2.4 à conclure que $M = L^*TL$ est un opérateur maximal monotone de K vers K. □

Si une des deux conditions de Gabay [8] est confirmée et si les images de L et de L^* sont fermées, alors le résultat annoncé par Gabay est vérifié.

En effet, si LL^* est un isomorphisme de H, alors $im\ L = H$ et donc L^*TL est maximal par le Théorème 3.2.6.

Si T^{-1} est fortement monotone et maximal, alors T doit être lipschitzien et maximal et donc en particulier localement borné sur H.
Nous utilisons le résultat suivant tiré de [4], que si nous avons un opérateur maximal monotone de H, alors l'opérateur est surjectif si et seulement si son inverse est localement borné. Ce qui implique que $D(T) = H$, et donc $im\ L = D(T)$ et de nouveau par le Théorème 3.2.6, nous obtenons que L^*TL est maximal.

Enfin, nous considérons le cas où $H = \mathbb{R}^n$. Nous passons donc de conditions sur les intérieurs à des conditions sur les intérieurs relatifs.

Rappelons, à titre indicatif, que nous avons un résultat dû à Minty [12] qui affirme que, si T est maximal monotone, alors son domaine $D(T)$ contient l'intérieur relatif (nécessairement non vide) de l'enveloppe convexe $conv\ D(T)$.

Proposition 3.2.7 *Soit T un opérateur maximal monotone de $\mathbb{R}^n$ vers $\mathbb{R}^n$ et soit N un sous-espace de $\mathbb{R}^n$. Si N rencontre $ri\ D(T)$, alors $T + \partial I_N$ est maximal monotone.*

Preuve: Voir [15].

Nous obtenons maintenant le résultat de maximalité en dimension finie, analogue au Théorème 3.2.6. L'hypothèse est affaiblie en demandant seulement une condition sur l'intérieur relatif, elle peut inclure la possibilité que $im\ L$ est l'espace $\mathbb{R}^n$ tout entier.

Théorème 3.2.8 *Supposons que L est une transformation linéaire de $\mathbb{R}^m$ vers $\mathbb{R}^n$. Soit T un opérateur maximal monotone de $\mathbb{R}^n$ vers $\mathbb{R}^n$. Si im L rencontre ri $D(T)$, alors $M = L^*TL$ est un opérateur maximal monotone de $\mathbb{R}^m$ vers $\mathbb{R}^m$.*

Preuve
Cette preuve découle des résultats précédents considérés en dimension finie. □

3.2.4 Conclusion

Dans ce chapitre, nous nous sommes intéressés à la composition de multifonctions. Nous avons prouvé un principe de dualité faisant intervenir dans les inclusions des compositions d'opérateurs maximaux monotones et de transformations linéaires. Ce principe nous a permis de réduire d'un point de vue théorique la difficulté de problèmes posés précédemment : ce principe diminue la dimension du problème initial.

Ensuite, nous nous sommes posés la question de savoir ce qu'il en était de la monotonie et la maximalité de cette composition puisqu'elle comportait un opérateur maximal monotone. Supposons que nous avons une inclusion comportant une somme d'opérateurs dont l'un est maximal monotone et l'autre est une composition comme donnée précédemment. Si nous prouvons que cette dernière est maximale monotone, nous pouvons appliquer la théorie développée à la fin du chapitre précédent pour montrer que la somme est maximale monotone.

Nous avons exposé une théorie qui améliore considérablement les connaissances précédentes, comme par exemple les conditions de Gabay [8].

Conclusion

Dans ce travail, nous avons développé deux principes de dualité, l'un pour une inclusion comportant une somme d'opérateurs, appelé principe de dualité abstraite dû à Attouch et Théra [1], et l'autre qui est l'extension du premier, pour une inclusion comportant une somme dont l'un des deux opérateurs est une composition d'opérateurs et d'applications linéaires. Nous avons, dans chacun des cas, explicité le principe et ensuite présenté leurs avantages, leurs inconvénients et surtout l'intérêt de leur application à un problème réel. Le premier principe de dualité nous a permis de démontrer un nouveau résultat quant à la monotonie et la maximalité d'une somme d'opérateurs, et le second de diminuer la dimension du problème initial pour résoudre des problèmes comportant une composition. Enfin, nous avons développé une théorie dûe à Robinson [15], qui nous donne une condition pour s'assurer de la monotonie et la maximalité de la composition. En perspective à ce travail, nous jugeons utile et bénéfique de considérer le côté pratique : une étude numérique consistant à résoudre des problèmes de grande taille (e.g., exemple économique), afin de montrer un des grands avantages de la composition.

Bibliographie

[1] H. Attouch and M. Théra, A general duality principle for the sum of two operators, *Journal of Convex Analysis*, 3 : 1-24, 1996.

[2] Jean-Pierre Aubin, *L'analyse non linéaire et ses motivations économiques*, Masson, Paris, 1984.

[3] Jean-Pierre Aubin, Ivar Ekeland, *Applied nonlinear analysis*, Wiley-Intersciencce, New York, 1984.

[4] H. Brézis, *Opérateurs maximaux monotones et semi-groupes de contraction dans les espaces de Hilbert*, Number 5 in North-Holland Mathematics Studies, North-Holland, Amsterdam, 1973

[5] H. Brézis, A. Haraux, Images d'une somme d'opérateurs monotones et applications, Israel J. Math 23, 2 : 165-186, 1976.

[6] Ivar Ekeland et R. Temam, *Analyse convexe et problèmes variationnels*, Dunod Gauthier-Villars, Paris, 1974.

[7] M. Fukushima, The primal Douglas-Rachford splitting algorithm for a class of monotone mappings with appplication to the traffic equilibrium problem, *Mathematical Programming*, 72 : 1-15, 1996.

[8] D. Gabay, Applications of the method of multipliers to variational inequalities, In M. Fortin and R. Glowinski editors, *Augmented Lagrangian Methods : Applications to the Numerical Solution of Boundary-Value Problems*, number 15 in Studies in Mathematics and Its Applications, page 299-331, North-Holland, Amsterdam, 1983.

[9] R. Glowinski, Jacques-Louis Lions et R. Trémolières, *Analyse numérique des inéquations variationnelles*, volume 1, Dunod, Paris, 1976.

[10] R. Glowinski, Jacques-Louis Lions et R. Trémolières, *Analyse numérique des inéquations variationnelles*, volume 2, Dunod, Paris, 1976.

[11] Pierre Jean Laurent, *Approximation et Optimisation*, Herman, Paris, 1972.

[12] G.J. Minty, On the maximal domain of a "monotone" function, *Michigan Mathematical Journal*, 8 : 135-137, 1966.

[13] U. Mosco, Dual variational inequalities, *Journal of Mathematical Analysis and Applications*, 40 : 202-206, 1972.

[14] Dominique Orban, *Elargissement d'opérateurs monotones - Dualité et inéquations variationnelles*, Mémoire, FUNDP, Namur, 1997.

[15] Stephen.M. Robinson, *Composition Duality and Maximal Monotonicity*, Departement of Industrial Engineering, University of Wisconsin-Madison, Madison, 1997.

[16] Stephen.M. Robinson, A reduction method for variational inequalities, *Mathematical Programming*, 79 : 101-109, 1997.

[17] R.Tyrrell. Rockafellar, *Convex analysis*, Princeton Univ. Press, 1970.

[18] R.Tyrrell. Rockafellar, On the maximality of sums of nonlinear monotone operators, *Transactions of the American Mathematical Society*, 149 : 75-88, 1970.

[19] R.Tyrrell. Rockafellar, *Conjugate duality and optimization*, Number 16 in CBMS Regional Conference Series in Applied Mathematics, Society for Industrial and Applied Mathematics, Philadelphia, PA, 1974.

[20] Michel Willem, *Analyse convexe et optimisation*, Cabay, Louvain-La-Neuve, 1986.